



UNIVERSIDADE FEDERAL DO PARÁ
CAMPUS UNIVERSITÁRIO DE CASTANHAL
FACULDADE DE MATEMÁTICA
CURSO DE LICENCIATURA PLENA EM MATEMÁTICA

JÚLIA RODRIGUES BARBOSA

**ANÁLISE ESTATÍSTICA E DE SÉRIES TEMPORAIS NA MODELAGEM
MATEMÁTICA DA INADIMPLÊNCIA NO SERASA**

CASTANHAL – PA
AGOSTO/2023

JÚLIA RODRIGUES BARBOSA

**ANÁLISE ESTATÍSTICA E DE SÉRIES TEMPORAIS NA MODELAGEM
MATEMÁTICA DA INADIMPLÊNCIA NO SERASA**

Monografia produzida como Trabalho de Conclusão de Curso, apresentado à comissão examinadora da Faculdade de Matemática do *Campus* de Castanhal da Universidade Federal do Pará, como requisito parcial para obtenção do grau de Licenciatura em Matemática, sob a orientação do Prof. Renato Germano Reis Nunes.

CASTANHAL – PA
AGOSTO/2023

**ANÁLISE ESTATÍSTICA E DE SÉRIES TEMPORAIS NA MODELAGEM
MATEMÁTICA DA INADIMPLÊNCIA NO SERASA**

Trabalho de Conclusão de Curso orientado pelo Prof. Renato Germano Reis Nunes, apresentado à Faculdade de Matemática como requisito para obtenção do grau de Licenciado em Matemática.

APROVADO EM: 30/08/2023

Conceito: EXCELENTE

BANCA EXAMINADORA:

Orientador: Prof. Dr. Renato Germano Reis Nunes/FACMAT/UFPA

Membro: Prof. Dr. Arthur Da Costa Almeida/FACMAT/UFPA

Membro: Prof. Dr. Edilberto Oliveira Rozal/FACMAT/UFPA

“Ensinar não é uma função vital,
porque não tem o fim em si mesma; a
função vital é aprender.”

(Aristóteles)

AGRADECIMENTOS

A Deus, que fez com que meus objetivos fossem alcançados, durante todos os meus anos de estudos e por me permitir ultrapassar todos os obstáculos encontrados ao longo da realização deste trabalho.

Aos meus pais, pelo apoio constante e pela dedicação. Pelo incentivo nos momentos difíceis, pela compreensão da minha ausência e, principalmente, por terem sido exemplos de vida, grandes referências profissionais e influenciadores significativos na escolha da docência, em especial, na área da matemática.

Aos meus irmãos, pela torcida, boas energias e vibração ao comemorar cada conquista no decorrer dessa caminhada, principalmente à Larissa, quem compartilhou todo o processo e foi uma das maiores incentivadoras para que eu concluísse essa etapa, mas, que acima de tudo, foi minha base e acalento nos momentos mais difíceis, quem compreendeu e quem estaria lá acaso algo fugisse do esperado.

Aos colegas de turma, pela amizade incondicional e pelo apoio demonstrado ao longo de todo o período. Em especial, aos amigos, Felipe Navarro, João Casseb, Luan Matheus e Valéria Borges, que me acolheram na turma e foram pessoas essenciais para que essa jornada se tornasse mais leve e prazerosa. Ao Luan Diego “da 2019”, quem esteve comigo assim que entrou no campus e que muito contribuiu no decorrer do curso, meu parceiro e grande amigo, o meu muito obrigada pela amizade e colaboração no decorrer desses anos.

Ao meu parceiro de vida, Tarciso, pela paciência e compreensão em momentos de incerteza e angústia sentidas ao longo do tempo.

Aos professores dessa instituição, pelas correções e ensinamentos que me permitiram apresentar um melhor desempenho no meu processo de formação profissional ao longo do curso.

Ao meu orientador, professor Renato Germano, meus agradecimentos em especial, pela contribuição, paciência e parceria no desenvolvimento deste trabalho de conclusão de curso, sem o qual nada disso seria possível.

À instituição de ensino UFPA – Campus Castanhal, essencial no meu processo de formação profissional, por tudo o que aprendi ao longo do curso.

A todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigada.

RESUMO

A presente pesquisa tem com o objetivo de apresentar uma modelagem matemática para a análise do índice de inadimplentes do Serasa, com ênfase na diferenciação por gênero e idade dos consumidores, como exemplo de aplicação para um curso de Licenciatura em Matemáticas. Utilizando dados históricos para construir modelos estatísticos destinados a descrever a evolução da inadimplência ao longo do tempo. Ao empregar técnicas de regressão e análise de variância, o estudo se esforça para descobrir padrões e tendências nos dados. Isso é obtido por meio de análises estatísticas e representações gráficas, aprimorando a clareza e a intuitividade da visualização dos dados. Essas descobertas denotam a presença de recorrências ou estruturas consistentes dentro de um conjunto de dados, além de oferecer percepções para tomada de decisão informada, previsão de comportamento futuro e compreensão de fatores de influência nos dados estudados. Deste modo, a aplicação da modelagem matemática ao referido objeto de estudo oferece aos alunos a oportunidade de mergulhar em conceitos matemáticos avançados dentro de um contexto pertinente, expandindo assim sua compreensão da utilidade da matemática nas ciências sociais e financeiras. Ao permitir que os alunos se envolvam com esses conceitos matemáticos avançados em uma estrutura do mundo real, suas habilidades analíticas e de resolução de problemas podem ser refinadas. Esta aplicação prática a um cenário relevante e tangível dentro da sociedade contemporânea nutre ainda mais suas capacidades.

Palavras-Chave: Séries temporais. Serasa. Modelagem Matemática. Inadimplência.

LISTA DE TABELAS

TABELA 1 - Resumo da estatística descritiva dos dados analisados.....	16
TABELA 2 - Resumo da estatística do teste t Student.....	16

LISTA DE GRÁFICOS

GRÁFICO 1 - Resultado do teste t, em que não há diferença entre as amostras.....	17
GRÁFICO 2 - Série histórica dos inadimplentes em função do sexo.....	18
GRÁFICO 3 - Série histórica dos inadimplentes em função da faixa etária.....	18
GRÁFICO 4 - Média móvel dos dados totais.....	19
GRÁFICO 5 - Tendência dos dados totais, juntamente com a série original.....	19
GRÁFICO 6 - Sazonalidade dos dados totais.....	20
GRÁFICO 7 - Decomposição residual dos dados totais.....	20
GRÁFICO 8 - Autocorrelação dos dados totais.....	20
GRÁFICO 9 - Decomposição espectral dos dados totais.....	21
GRÁFICO 10 - Previsão do modelo ARIMA (1, 1, 1) x (1, 1, 1, 12) [12]	21

SUMÁRIO

1.INTRODUÇÃO	10
2.METODOLOGIA	11
3.CONCEITOS FUNDAMENTAIS DA ANÁLISE ESTATÍSTICA	12
3.1.Média móvel	13
3.2.Decomposição.....	13
3.2.1.Tendência.....	13
3.2.2.Sazonalidade	14
3.2.3.Componente Residual	14
3.3.Autocorrelação	14
3.4.Decomposição espectral	14
3.5.Modelo ARIMA	15
4.FATORES E VARIÁVEIS DETERMINANTES DA INADIMPLÊNCIA	16
5.RESULTADOS E DISCUSSÃO	16
5.1.Por Sexo.....	18
5.2.Por Idade	18
5.3.Média móvel	19
5.4.Decomposição.....	20
5.4.1.Tendência.....	20
5.4.2.Sazonalidade	20
5.4.3.Componente Residual	20
5.5.Autocorrelação	21
5.6.Decomposição espectral	21
5.7.Modelo ARIMA	21
6.CONSIDERAÇÕES FINAIS	23
7.DIREÇÕES FUTURAS DE PESQUISA	24
REFERÊNCIAS	25
APÊNDICE	26

1. INTRODUÇÃO

A Modelagem Matemática é uma poderosa ferramenta que permite aos matemáticos analisarem fenômenos complexos do mundo real. A importância desta ferramenta tem reflexo no processo educacional, desde o Ensino Básico até o Ensino Superior. Nesse sentido, no caso da Base Nacional Comum Curricular – BNCC, é possível encontrar diversas referências ao processo de modelagem matemática no contexto da educação básica, tais como: “modelagem”, “construção de modelos”, inclusive nas competências específicas de matemática para o Ensino Fundamental e o Ensino Médio (BRASIL, 2018). Assim, diante disso, os alunos dos cursos de licenciatura em Matemática, têm que, necessariamente, adquirir uma base de como desenvolver modelos e interpretá-los.

Dentre as mais diversas técnicas matemática disponíveis para o desenvolvimento de modelos matemáticos, a análise estatística é uma abordagem quantitativa que envolve a coleta, organização, interpretação e apresentação de dados para obter percepções e informações. Ela é utilizada para compreender padrões, relações, variações e tendências nos dados, possibilitando a tomada de decisões controladas. Envolve o uso de métodos estatísticos, como média, mediana, desvio padrão, regressão, análise de variância, entre outros, para resumir e analisar os dados, identificar relações entre variáveis e realizar inferências sobre uma população maior com base em uma amostra.

Nesse contexto, temos as séries temporais que são sequências de observações ou medidas registradas ao longo do tempo, em intervalos regulares ou irregulares. A análise de séries temporais é utilizada para estudar padrões e tendências que podem surgir ao longo do tempo, permitindo a identificação de comportamentos sazonais, cíclicos e outras variações temporais. Isso é especialmente útil em áreas como economia, finanças, meteorologia, epidemiologia e muitas outras disciplinas, onde o tempo desempenha um papel fundamental na compreensão e previsão de eventos (MORETTIN, 2006).

No contexto econômico e financeiro, a análise de dados temporais desempenha um papel crucial na identificação de tendências, padrões sazonais e flutuações irregulares. Este trabalho de conclusão de curso busca explorar a dinâmica subjacente aos comportamentos de consumidores inadimplentes, empregando técnicas de estatística descritiva e de modelagem de séries temporais. A aplicação de técnicas estatísticas e ferramentas numéricas de análise de séries temporais

permitirá discernir padrões subjacentes, identificar fatores determinantes e realizar previsões (BOX, JENKINS, REINSEL, LJUNG, 2016).

A inadimplência é caracterizada pelo não cumprimento de obrigações financeiras dentro dos prazos estabelecidos, é um fenômeno de considerável relevância em sistemas econômicos. Compreender os padrões temporais associados a esse comportamento é de suma importância para instituições financeiras, órgãos regulatórios e tomadores de decisão, uma vez que a inadimplência impacta diretamente a estabilidade financeira e a avaliação de riscos.

Assim, o presente estudo se propõe a investigar a evolução temporal da inadimplência por meio da análise de um conjunto de dados obtidos junto ao Serasa (SERASA, 2023). O Serasa é uma das maiores empresas de análise e informações de crédito no Brasil. O banco de dados do Serasa, nos permite coletar informações e dados históricos necessários para construir modelos estatísticos que descrevam o comportamento da inadimplência ao longo do tempo. O índice de inadimplência Serasa pessoa física é o principal indicador de inadimplência no país e é amplamente adotado por instituições públicas e privada, foi estabelecido através da integração de diversos índices de inadimplência. Essa unificação permitiu uma análise mais sólida e uniforme da inadimplência no Brasil, tornando possível realizar comparações e projeções com outros índices da economia brasileira.

Diante do exposto, o presente trabalho apresenta análises estatísticas e modelagem de séries temporais por meio de modelos ARIMA (Auto Regressive Integrated Moving Average), utilizando dados da quantidade de consumidores inadimplentes de acordo com a empresa Serasa, no período de março de 2016 até abril de 2023, como um exemplo de modelagem matemática no contexto de um curso de licenciatura em matemática.

2. METODOLOGIA

Esta pesquisa é de natureza básica, descritiva, cujo procedimento é um estudo de caso e com o método quantitativo como descrito por Prodanov e Freitas (2013).

Os dados históricos sobre o índice de inadimplência no Serasa foram coletados e organizados em uma série temporal. Para a modelagem matemática, utilizamos técnicas de regressão, considerando o índice de inadimplência como variável resposta e as variáveis gênero e idade dos consumidores como variáveis explicativas. Além disso, empregamos análise do teste t-student a fim investigar possíveis diferenças

significativas entre grupos.

Essas análises foram realizadas em Python (ATWAN, 2022; HUANG, PETUKHINA, 2022) no Ambiente de Desenvolvimento Integrado Jupyter Notebook (JUPYTER, 2023) o qual possui diversos repositórios de pacotes disponíveis.

A primeira etapa da metodologia se concentrou na coleta de dados histórico do índice de inadimplentes no Brasil, no site SerasaExperian, no qual a série temporal abrangia o período de março de 2016 até abril de 2023.

Em um segundo momento, ocorreu uma etapa de pré-processamento. Durante esta etapa, os dados coletados foram “limpos” para garantir qualidade e consistência, incluindo imputação de valores ausentes, correção de erros de digitação e padronização dos dados. Então, além de estimar estatísticas descritivas como média, mediana, desvio padrão e percentis, a análise exploratória é realizada usando gráficos de linhas para extrair as características iniciais da série temporal.

Através da decomposição sazonal, identificaram-se e isolaram-se os componentes de tendência, sazonalidade e resíduo na série temporal. Posteriormente, procedeu-se à modelagem dos componentes residuais utilizando-se a metodologia dos modelos ARIMA. A escolha do modelo ARIMA apropriado foi conduzida por meio de análise de autocorrelação, testes de estacionariedade e critérios de seleção de modelos, incluindo o critério de informação de Akaike (AIC).

Uma vez ajustado o modelo ARIMA aos dados, realizou-se a validação do mesmo para verificar sua capacidade preditiva. Tal procedimento foi realizado por meio de uma divisão dos dados em conjuntos de treinamento e teste, ajustando-se o modelo aos dados de treinamento e confrontando as previsões com os dados de teste. Com o modelo ARIMA devidamente validado, efetuaram-se projeções do índice de inadimplência para além do intervalo temporal disponível, fornecendo uma valiosa ferramenta para a estimativa de padrões futuros. Os resultados obtidos revelam a efetividade da abordagem estatística e de modelagem adotada neste estudo.

3. CONCEITOS FUNDAMENTAIS DA ANÁLISE ESTATÍSTICA

A análise estatística é um conjunto de técnicas e métodos utilizados para organizar, resumir, interpretar e tirar conclusões a partir de dados. Ela desempenha um papel fundamental em diversas áreas, como ciência, negócios, medicina, engenharia e muitas outras, ajudando a transformar dados brutos em informações

úteis e significativas.

As ferramentas estatísticas permitem resumir grandes volumes de dados de maneira concisa e compreensível. Isso ajuda a identificar tendências, padrões e características importantes nos dados. Ao analisar estatisticamente os dados, as decisões podem ser tomadas com base em evidências sólidas e não apenas em suposições ou intuições. Isso é crucial em negócios, pesquisa científica e governo. Através dessas análises é possível explorar relações de causa e efeito entre variáveis. Isso é essencial para entender como os fatores afetam um resultado específico e podem ser usados para melhorar processos.

Desta forma, este capítulo contempla os conceitos fundamentais das ferramentas estatísticas utilizadas para realizar as análises dos dados coletados.

3.1. Média móvel

A média móvel é um conceito estatístico e matemático amplamente utilizado em análise de séries temporais, estatísticas financeiras, processamento de sinais e muitas outras áreas. Ela é usada para suavizar flutuações em dados ao longo do tempo, tornando mais fácil identificar tendências, padrões subjacentes ou variações de curto prazo. A média móvel é calculada através da média de um determinado número de observações consecutivas em função do avanço do tempo.

3.2. Decomposição

A decomposição em série temporal é uma técnica analítica utilizada para entender os diferentes componentes que afetam a variação de uma série temporal ao longo do tempo. Ela é frequentemente empregada para entender tendências, padrões sazonais e variações residuais em dados sequenciais ao longo do tempo. A ajuda a separar esses componentes, permitindo uma compreensão mais clara do comportamento dos dados.

Uma série temporal geralmente pode ser composta em três componentes principais:

3.2.1. Tendência

Refere-se à direção geral dos dados ao longo do tempo. Pode ser ascendente (crescente), descendente (decrecente) ou horizontal (constante). A tendência de

capturar as mudanças de longo prazo nos dados.

3.2.2. Sazonalidade

A sazonalidade refere-se a padrões ou variações regulares e previsíveis que ocorrem em um conjunto de dados em intervalos específicos de tempo e são importantes para a análise e previsão de tendências. Ela é um componente comum em séries temporais e pode ser representada de diferentes maneiras, como picos e vales recorrentes em um gráfico ao longo do tempo.

3.2.3. Componente Residual

Também conhecido como erro, é a parte da série temporal que não pode ser explicada pela tendência ou pela sazonalidade. Inclui variações investigadas, influências externas imprevisíveis e outros fatores que não se enquadram nos padrões previsíveis.

3.3. Autocorrelação

A autocorrelação em séries temporais é uma medida estatística que avalia a relação entre os valores passados e presentes em uma série temporal. Em outras palavras, ela mede o grau de esplendor entre os pontos de dados em diferentes pontos no tempo. A autocorrelação é uma ferramenta importante para analisar a dependência serial dos dados em séries temporais, o que significa que valores passados podem influenciar os valores futuros. Isso é crucial para entender padrões e estruturas básicas nos dados, especialmente em especial em séries temporais.

A autocorrelação é utilizada para analisar se existe alguma estrutura de dependência entre os valores ao longo do tempo. O coeficiente de autocorrelação varia de -1 a 1, em que o valor de 1 indica uma correlação perfeita positiva entre os valores, -1 indica uma correlação perfeita negativa e zero indica ausência de correlação).

3.4. Decomposição espectral

A colocação espectral, também conhecida como análise de frequência, é uma técnica utilizada para entender os diferentes componentes de uma série temporal ou sinal em termos de suas frequências constituintes. É uma ferramenta útil para analisar padrões oscilatórios e sazonais em dados. A colocação espectral é frequentemente

aplicada através da Transformada de Fourier, que converte um sinal de domínio temporal em seu equivalente no domínio da frequência. A análise de Fourier permite identificar as frequências dominantes presentes em um sinal e separar seus componentes de acordo com suas frequências.

3.4.1. Modelo ARIMA

Durante a análise dos dados, empregou-se modelos matemáticos com habilidade para gerar compreensão. Nesse sentido, optou-se pelo uso do modelo estatístico ARIMA, que apresenta uma ampla aplicação para realizar previsões de séries temporais, sendo considerado um dos métodos mais eficazes disponíveis.

ARIMA, que significa *AutoRegressive Integrated Moving Average*, é um modelo estatístico usado para previsão de séries temporais. Ele combina três fatores em sua análise: O componentes autorregressivos (AR), nesse componente, os valores anteriores da série temporal são usados para prever o valor futuro. Isso implica que as observações passadas têm influência sobre o futuro. O componente de média móvel (MA), onde os erros passados (ou resíduos) da previsão são usados para prever o valor futuro. Isso significa que os erros passados contemplaram as perspectivas futuras. E diferenciação integrada (I) para modelar padrões temporais complexos em dados, refere-se à transformação da série temporal para fazer a estacionária, ou seja, remover tendências e sazonalidades que dificultam a análise. O termo "integrado" refere-se à aplicação de operações de diferenciação para atingir a estacionariedade.

A estrutura básica de um modelo ARIMA é definida pelas seguintes parâmetros: Ordem Autoregressiva (p): Refere-se ao número de valores anteriores da série temporal que são usados para prever o próximo valor. Quanto maior o valor de "p", mais a previsão depende de valores passados. Ordem de Diferenciação (d): É o número de vezes que a série temporal é diferenciada para a seleção estacionária. A diferenciação é feita para remover tendências e sazonalidades, tornando os dados mais adequados para modelos. Ordem Média Móvel (q): Representa o número de erros passados que são usados para prever o próximo valor. Da mesma forma que não há componente autorregressivo, quanto maior o valor de "q", mas a previsão depende de erros passados.

4. FATORES E VARIÁVEIS INFLUENTES DA INADIMPLÊNCIA

A inadimplência, ou seja, a falta de pagamento de dívidas ou obrigações financeiras, é influenciada por uma série de fatores e variáveis que podem variar dependendo do contexto e das características do devedor.

Alguns dos principais fatores e variáveis que influenciam a inadimplência são: histórico de crédito, renda e capacidade disponível, nível de endividamento, taxas de juros, prazo e valor da dívida, bem como o consumismo, são fatores gerados pela carência de educação financeira, caracterizada pela falta de conhecimento e habilidades para tomar decisões financeiras informadas e responsáveis. Assim, torna-se um problema comum em muitas partes do mundo e pode ter várias consequências negativas para indivíduos e sociedades.

O histórico de crédito de um indivíduo é um dos principais fatores que influenciam a inadimplência. Se alguém tem um histórico de pagamento pontual e responsável, é menos provável que se torne inadimplente. Por outro lado, um histórico de atrasos e falta de pagamento aumenta o risco.

O endividamento está em ascensão globalmente, e embora não seja um desafio exclusivo do Brasil, merece atenção especial devido à sua extensão, conforme mencionado por Ferreira (2009). Essa tendência pode ser atribuída à falta de conhecimento e planejamento financeiro, fatores pessoais influenciados por aspectos sociais e psicológicos, como descritos nas finanças comportamentais, bem como as influências externas, como desemprego, redução de renda, questões de saúde e outros fatores.

De acordo com uma pesquisa realizada pelo Serasa, em parceria com o instituto Opinion Box, empresa de tecnologia referência em pesquisa de mercado e customer experience, realizada na 5ª edição do levantamento anual sobre o cenário do endividamento no Brasil, através de entrevistas online com 5.225 consumidores da base de endividados do Serasa, sendo 50% homens e 50% mulheres, o desemprego segue como o principal motivos do endividamento em 2022. Mesmo seguindo uma tendência de queda, o desemprego ainda é a principal causa para o endividamento. Ele impacta principalmente mulheres, representando 31% e jovens de até 30 anos, representando 33% dos endividados. Outros motivos do endividamento destacado na pesquisa entre os brasileiros são redução de renda, representando 12%, alguém que compra em seu nome e não fez o pagamento, 8% e, a falta de controle representando 16%.

5. RESULTADOS E DISCUSSÃO

Da análise de conteúdo, realizada com base nos dados recolhidos, importa salientar o quanto o discente enquanto pesquisador é colocado em posição oportuna para aplicar o conteúdo e conhecimento obtido em sala de aula no decorrer da graduação. No presente estudo, podemos utilizar a modelagem matemática dentro do conceito da disciplina de Probabilidade e Estatística, permitindo a interdisciplinaridade. Visto que, a modelagem matemática envolve a criação e uso de representações matemáticas para descrever, analisar e compreender conceitos da vida real. Ela é uma abordagem que busca traduzir situações complexas e problemas do mundo real em termos matemáticos, permitindo assim que sejam treinados e resolvidos com ferramentas matemáticas.

Por tanto, através dos dados de índice de inadimplência do Serasa, obtivemos a oportunidade de estudar, de forma exploratória, um contexto real das ciências sociais e financeiras. De modo que a manipulação desses dados nos permite analisar e estudar os possíveis impactos, consequências e desequilíbrio financeiro pessoal, além de poder ajudar a adotar estratégias de gerenciamento de risco para minimizar a exposição à inadimplência.

A modelagem matemática da inadimplência no Brasil, assim como em qualquer lugar, é um desafio complexo, pois envolve muitas variáveis interconectadas e comportamentos humanos imprevisíveis. No entanto, algumas abordagens estatísticas e matemáticas podem ser usadas para tentar entender e prevenir a inadimplência. A análise de regressão, por exemplo, pode ajudar a identificar quais variáveis têm uma clareza mais forte com a inadimplência e como elas interagem entre si.

Além disso, técnicas mais avançadas, como modelos de séries temporais, pode ser aplicada para melhorar a precisão das direções. Esses modelos podem levar em consideração não apenas variáveis econômicas, mas também dados comportamentais, como histórico de pagamento, padrões de gastos e outros fatores que podem indicar uma probabilidade de inadimplência.

Dessa forma, a discussão dos resultados permitirá que os estudantes compreendam como a modelagem matemática pode ser aplicada na interpretação de dados financeiros complexos, possibilitando a identificação de fatores que influenciam a inadimplência no

Serasa. Além disso, os alunos terão a oportunidade de discutir as limitações dos

modelos utilizados e a importância de considerar outros fatores que podem afetar a inadimplência.

A estatística descritiva dos dados analisados é mostrada na Tabela 1.

Tabela 1: Resumo da estatística descritiva dos dados analisados.

	N	Omisso	Média	Mediana	Desvio-padrão	Mínimo	Máximo
Consumidores Inadimplentes (milhões)	86	0	62,82	62,49	3,204	58,63	71,44
Dívida Média (R\$)	86	0	3982,58	3927,66	214,028	3624,78	4767,20
Masculino	52	34	32,24	31,50	1,506	30,59	35,93
Feminino	52	34	32,44	32,21	1,259	30,78	35,51
Até 25 anos (milhões)	52	34	8,34	8,50	0,483	7,48	9,14
26 - 40 anos	52	34	23,38	23,51	0,810	22,03	25,16
41 - 60 anos	52	34	22,26	21,93	0,885	20,94	24,49
Acima de 60 anos	52	34	10,71	10,55	0,936	9,19	12,65

Fonte: própria pesquisa.

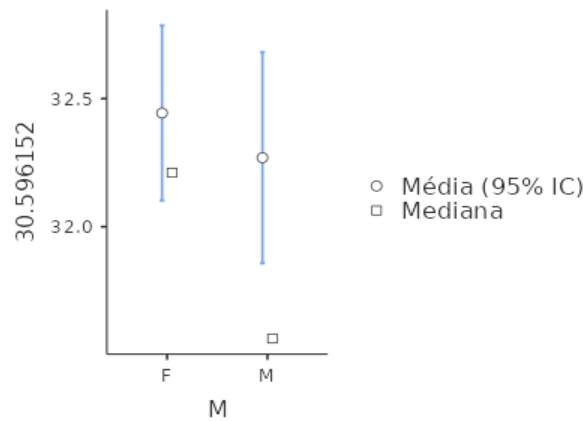
Com o objetivo de identificar fatores que influenciam o índice de inadimplência, levamos em consideração a variável sexo, investigando se há alguma diferença dos grupos masculino e feminino realizando o teste t para amostras independentes, cujo resultado pode ser visto no Gráfico 1.

Tabela 2: Resumo da estatística do teste t Student.

	Estatística	gl	p	Diferença média	Erro-padrão da Diferença
t de Student	0,639	101	0,525	0,174	0,273

Fonte: própria pesquisa.

Gráfico 1: Resultado do teste t, em que não há diferença entre as amostras.



Fonte: própria pesquisa.

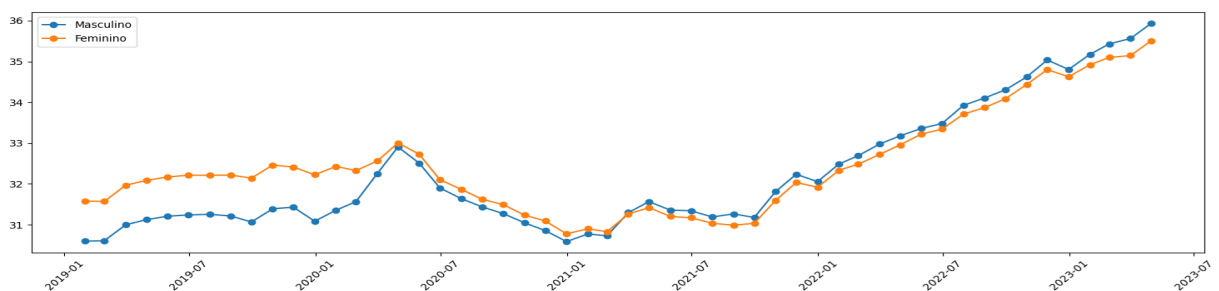
Assim, é possível notar que não há diferenças significativas entre os grupos, o que nos leva a concluir que no quesito inadimplência, homens e mulheres se comportam, estatisticamente, iguais.

A seguir apresentamos as séries temporais dos inadimplentes e suas principais análises.

5.1. Por sexo

No Gráfico 2 vemos o número de consumidores inadimplentes (em milhões) no intervalo de 2019 até 2023, em função do sexo.

Gráfico 2: Série histórica dos inadimplentes em função do sexo (valores em milhões).

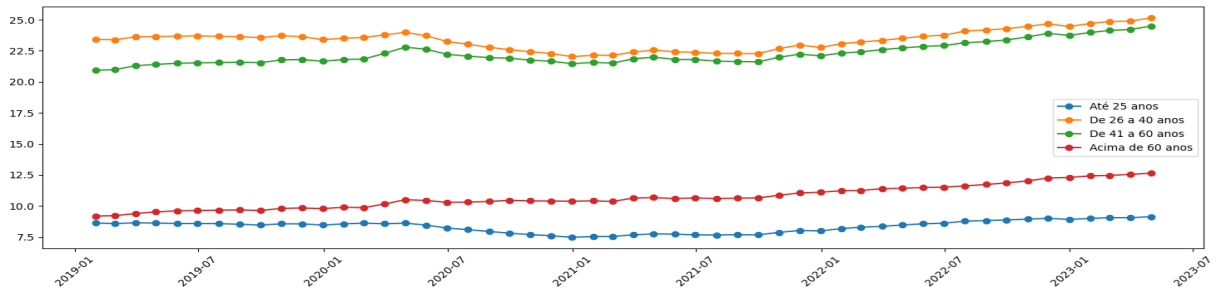


Fonte: própria pesquisa.

5.2. Por idade

No Gráfico 3 vemos o número de consumidores inadimplentes (em milhões) no intervalo de 2019 até 2023, para algumas faixas etárias.

Gráfico 3: Série histórica dos inadimplentes em função da faixa etária (valores em milhões).



Fonte: própria pesquisa.

As faixas etárias com maior índice de inadimplência podem variar de acordo com fatores econômicos, sociais e demográficos, mas historicamente algumas faixas etárias apresentam maiores desafios financeiros.

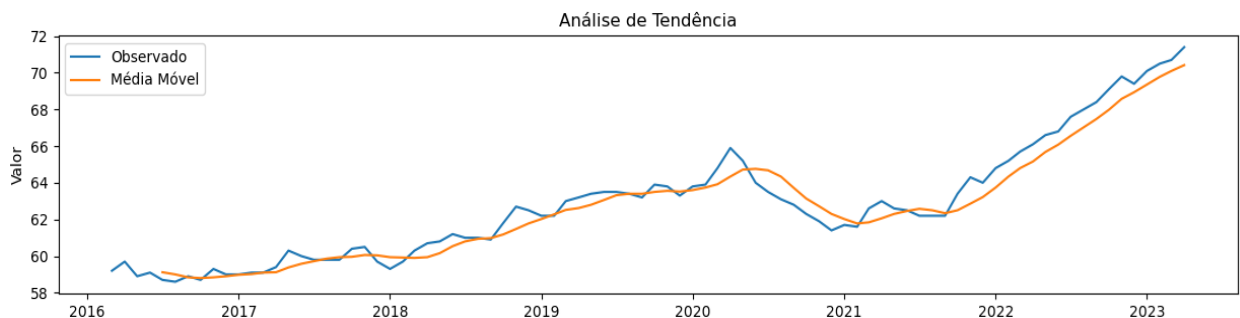
De acordo com o Gráfico 3, é possível notar que a faixa com maior índice de inadimplência é de 26 até 40 anos. Pessoas nessa faixa etária muitas vezes estão ingressando no mercado de trabalho, lidando com as responsabilidades financeiras pela primeira vez e não podem ter uma educação financeira sólida. Isso pode levar a decisões financeiras equivocadas, como aquisição de dívidas sem um plano sólido de pagamento.

Em seguida temos a faixa de 41 a 60 anos como segundo maior índice. Essa faixa etária pode enfrentar desafios relacionados a gastos familiares, educação dos filhos, saúde e, às vezes, uma maior pressão para manter um padrão de vida. Esses fatores podem aumentar o risco de inadimplência.

5.3. Média móvel

O Gráfico 4 mostra os dados observados e a Média Móvel do índice de inadimplentes.

Gráfico 4: Média móvel dos dados totais.

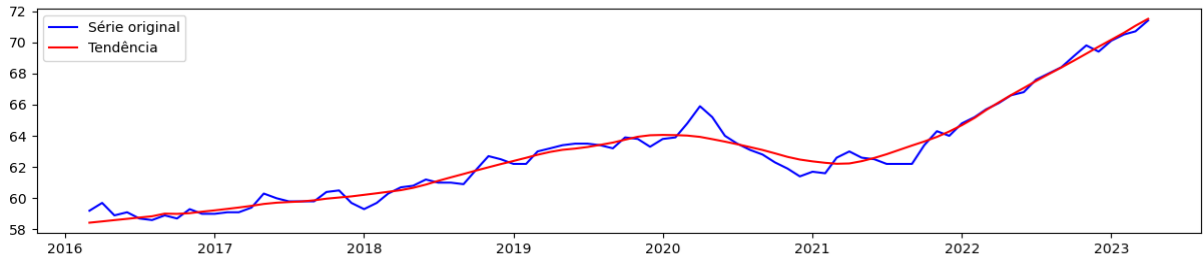


Fonte: própria pesquisa.

5.4. Decomposição

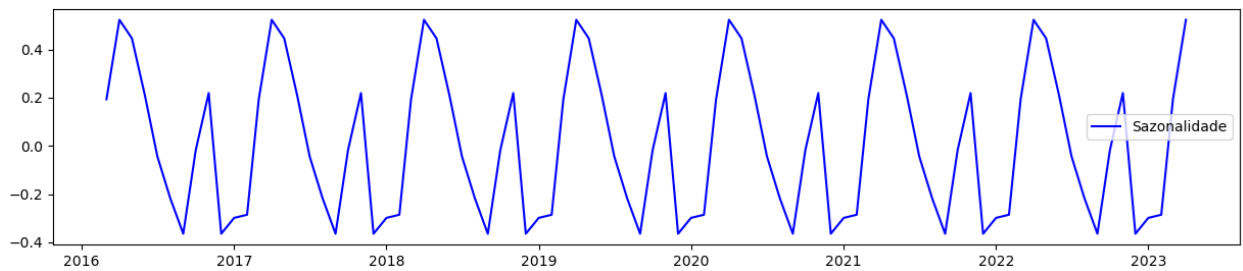
5.4.1. Tendência

Gráfico 5: Tendência dos dados totais, juntamente com a série original.



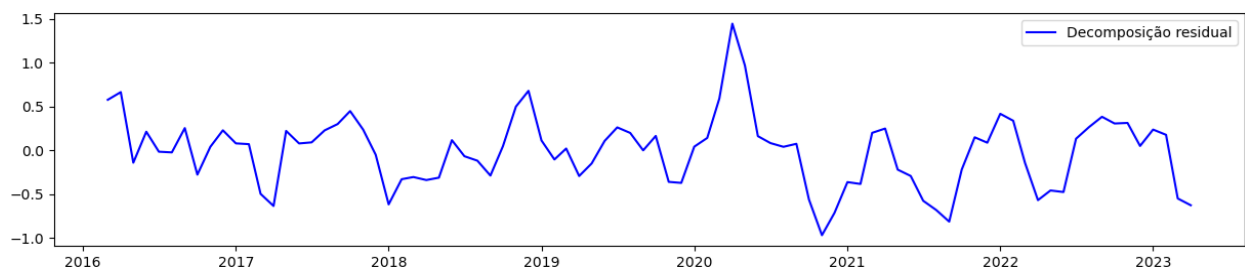
5.4.2. Sazonalidade

Gráfico 6: Sazonalidade dos dados totais.



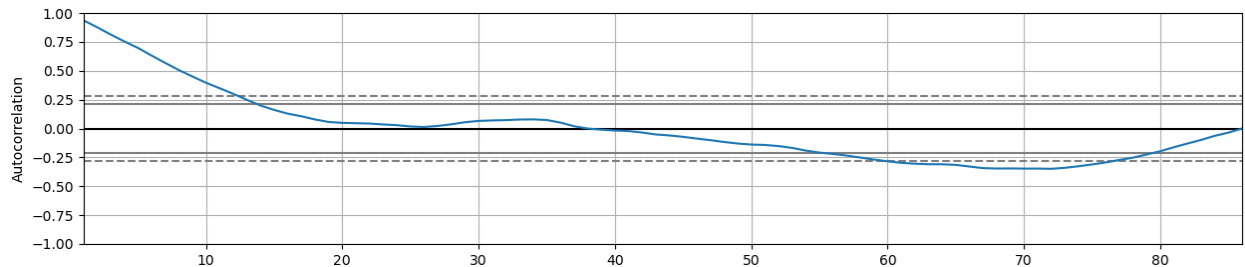
5.4.3. Componente Residual

Gráfico 7: Decomposição residual dos dados totais.



5.5. Autocorrelação

Gráfico 8: Autocorrelação dos dados totais.

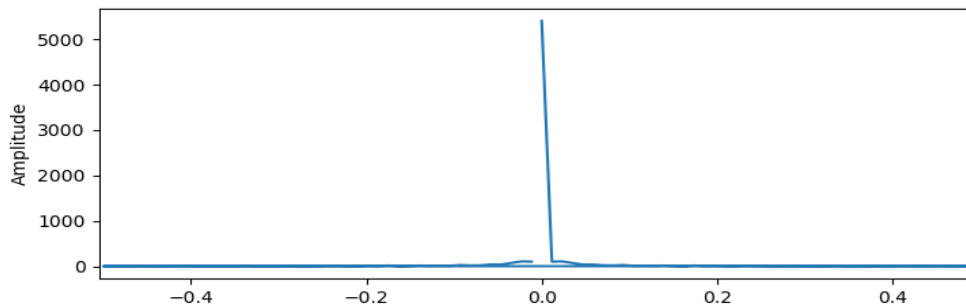


Fonte: própria pesquisa.

5.6. Decomposição espectral

Utilizamos a chamada transformada rápida de Fourier (FFT) cujos resultados são mostrados no Gráfico 9.

Gráfico 9: Decomposição espectral dos dados totais.



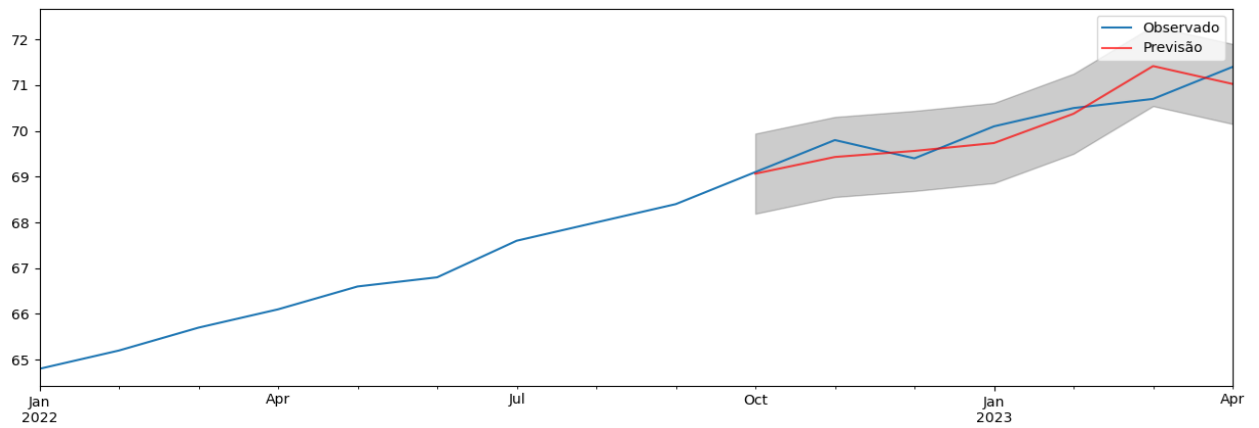
Fonte: própria pesquisa.

É possível notar que não existe um valor de frequência dominante na série temporal, a qual indica que os dados não apresentam uma frequência característica regular.

5.7. Modelo ARIMA

Realizada a adequação, o modelo ARIMA foi utilizado no período de janeiro à setembro de 2022, e previsto entre os períodos de outubro de 2022 à abril de 2023, a qual foi mostrado no Gráfico 10.

Gráfico 10: Previsão do modelo ARIMA(1, 1, 1)x(1, 1, 1, 12)[12].



Fonte: própria pesquisa.

É possível notar que, a partir de outubro de 2022, a previsão é bastante satisfatória o que demonstra que os métodos utilizados em nossa análise são adequados para previsões futuras do índice de inadimplentes.

6. CONSIDERAÇÕES FINAIS

Este exemplo de aplicação ressalta a importância da matemática como uma ferramenta de grande relevância na avaliação de dados tanto financeiros quanto sociais. O uso da modelagem matemática para analisar a inadimplência registrada no Serasa oferece aos estudantes a oportunidade de explorar conceitos avançados, como regressão, séries temporais e análise de variância, inseridos em um contexto real. A implementação prática da matemática em situações do mundo concreto, como esta em questão, contribui para uma formação abrangente e enriquecedora. Isso prepara os futuros educadores em matemática para enfrentar desafios intrincados, ao mesmo tempo que promove uma compreensão mais profunda da relevância da matemática nas esferas sociais e financeiras.

Torna-se evidente a relevância significativa desse estudo no cenário financeiro e econômico. Uma análise aprofundada dos índices de inadimplência oferece percepções valiosas sobre as dinâmicas subjacentes que afetam a estabilidade financeira de indivíduos, empresas e da economia em geral.

Os modelos matemáticos e as abordagens estatísticas aplicadas forneceram uma visão mais clara dos padrões de comportamento e das tendências ao longo do tempo, permitindo uma melhor antecipação dos riscos associados.

É crucial destacar que a modelagem matemática desses índices não só contribui para a compreensão teórica, mas também para a aplicação prática. Instituições financeiras, empresas e tomadores de decisão podem utilizar os resultados dessa pesquisa para desenvolver estratégias de gerenciamento de riscos, concessão de crédito e tomada de decisões financeiras informadas.

No entanto, é importante considerar as limitações do estudo, como a dependência de dados históricos e a sensibilidade a mudanças abruptas no cenário econômico. Portanto, sugere-se que pesquisas futuras considerem uma análise mais ampla e aprofundada, incorporando dados em tempo real e considerando cenários diversos. Com o intuito de ressaltar a importância de uma abordagem multidisciplinar ao analisar os índices de inadimplência, considerando não apenas fatores financeiros, mas também sociais, econômicos e psicológicos que podem influenciar o comportamento de pagamento.

Em resumo, uma pesquisa sobre os índices de inadimplência não apenas enriqueceu nosso entendimento das complexidades financeiras, mas também oferece

exemplos práticos para enfrentar desafios econômicos e financeiros de maneira mais eficaz e informada.

7. DIREÇÕES FUTURAS DE PESQUISA

Apesar das contribuições significativas deste estudo, existem áreas adicionais que merecem exploração futura:

Modelos Avançados: A investigação de modelos mais complexos, como modelos sazonais (SARIMA) e modelos de séries temporais exponenciais (ETS), pode aprimorar ainda mais a precisão das previsões.

Incorporação de Dados Externos: A inclusão de variáveis externas, como indicadores econômicos e dados demográficos, pode enriquecer os modelos de previsão e oferecer insights adicionais sobre os determinantes da inadimplência.

Análise Espacial: Explorar padrões geográficos da inadimplência, combinando dados temporais com informações de localização, pode fornecer *insights* sobre variações regionais no comportamento de inadimplência.

Análise de Causa: Investigações futuras podem se concentrar na identificação das causas subjacentes à inadimplência, através da combinação de dados financeiros e informações socioeconômicas.

REFERÊNCIAS

ATWAN, T. A., Time Series Analysis with Python Cookbook: Practical recipes for exploratory data analysis, data preparation, forecasting, and model evaluation, Packt, Birmingham, 2022.

BRASIL. Ministério da Educação. Base Nacional Comum Curricular. Brasília, 2018.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M., Time Series Analysis: Forecasting and Control, 5a ed., John Wiley & Sons, Inc., Hoboken, New Jersey, 2016.

CRYER, J. D.; CHAN, K.-S., Time Series Analysis: With Applications in R, 2a ed., Springer, New York, 2008.

HUANG, C.; PETUKHINA, A., Applied Time Series Analysis and Forecasting with Python, Springer, Cham, 2022.

JUPYTER, <https://jupyter.org/>, acesso em agos. de 2023.

MORETTIN, P. A., Análise de séries temporais, 2a ed., Blucher, São Paulo, 2006.

PRODANOV, C. C.; FREITAS, E. C. Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico, 2a ed., Universidade Feevale, Novo Hamburgo, 2013.

SERASA, <https://www.serasa.com.br/>, acesso em ago. de 2023.

APÊNDICE

Desta monografia foi gerado um texto que foi aceito para publicação como capítulo de livro na Editora Científica, intitulado “MODELAGEM MATEMÁTICA DA INADIMPLÊNCIA NO SERASA: UMA ANÁLISE ESTATÍSTICA E DE SÉRIE TEMPORAL”, o qual pode ser visto a seguir, juntamente com a carta de aceite para publicação.

MODELAGEM MATEMÁTICA DA INADIMPLÊNCIA NO SERASA: UMA ANÁLISE ESTATÍSTICA E DE SÉRIE TEMPORAL

Júlia Rodrigues Barbosa

Faculdade de Matemática da Universidade Federal do Pará, Campus Castanhal

Renato Germano

Faculdade de Matemática da Universidade Federal do Pará, Campus Castanhal

RESUMO

Objetivo: Este artigo apresenta uma modelagem matemática para a análise do índice de inadimplentes no Serasa, com ênfase na diferenciação por gênero e idade dos consumidores, como exemplo de aplicação para um curso de Licenciatura em Matemática. A pesquisa utiliza dados históricos para construir modelos estatísticos que descrevam o comportamento da inadimplência ao longo do tempo. Utilizando técnicas de regressão e análise de variância, o estudo busca identificar padrões e tendências nos dados através de análises estatísticas e representações gráficas que permitem visualizar os dados de forma mais clara e intuitiva. A análise de padrões e tendências é uma parte essencial da análise exploratória de dados e da modelagem estatística. Essas informações indicam a presença de repetições ou estruturas consistentes dentro de um conjunto de informações, além de poder ajudar a tomar decisões informadas, prever comportamentos futuros e entender os fatores que influenciam os dados em estudo. A aplicação da modelagem matemática permite aos estudantes explorar conceitos avançados de matemática em um contexto real e relevante, ampliando sua compreensão da utilidade da matemática nas ciências sociais e financeiras. Ao permitir que os estudantes explorem conceitos avançados de matemática em um contexto real, eles podem aprimorar suas habilidades analíticas e de resolução de problemas, aplicando-as a uma situação concreta e pertinente no campo da sociedade contemporânea. Além disso, essa abordagem multidisciplinar demonstra aos futuros professores de matemática a ampla gama de aplicações da disciplina além dos tradicionais problemas matemáticos, reforçando a importância e a utilidade da matemática no mundo real.

Palavras-chave: Séries temporais. Serasa. Modelos ARIMA. Inadimplência.

INTRODUÇÃO

A modelagem matemática é uma poderosa ferramenta que permite aos matemáticos analisarem fenômenos complexos do mundo real. A importância desta ferramenta tem reflexo no processo educacional, desde o Ensino Básico até o Ensino Superior. Nesse sentido, no caso da Base Nacional Comum Curricular – BNCC, é possível encontrar diversas referências ao processo de modelagem matemática no contexto da educação básica, tais como: “modelagem”, “construção de modelos”, inclusive nas competências específicas de matemática para o Ensino Fundamental e o Ensino Médio (BRASIL, 2018). Assim, diante disso, os alunos dos cursos de licenciatura em Matemática, têm que, necessariamente, adquirir uma base de como desenvolver modelos e interpretá-los. Desta forma, apresentaremos uma análise matemática (estatística) da quantidade de consumidores inadimplentes de acordo com a empresa Serasa, no período de março de 2016 até abril de 2023, como um exemplo de modelagem matemática no contexto de um curso de licenciatura em matemática.

Inadimplência é um termo usado para descrever o fenômeno financeiro em que uma pessoa ou entidade deixa de cumprir suas obrigações financeiras, como o pagamento de dívidas, empréstimos, contas ou outras responsabilidades financeiras. Quando alguém está inadimplente, significa que está atrasado ou não pagou o valor devido dentro do prazo acordado, o que pode resultar em consequências negativas tanto para os devedores quanto para os credores, e requerendo ações cautelosas para evitar e gerenciar essa situação.

Um dos principais indicadores econômicos dentro do banco de dados do Serasa é o índice de inadimplência do consumidor pessoa física, amplamente adotado por instituições públicas e privada, foi estabelecido através da integração de diversos índices de inadimplência. Essa unificação permitiu uma análise mais sólida e uniforme da inadimplência no Brasil, tornando possível realizar comparações e projeções com outros índices da economia brasileira.

Neste trabalho, estudaremos a quantidades de consumidores inadimplentes, buscamos descrever e interpretar padrões históricos de inadimplência através da análise estatística e modelagem de séries temporais, considerando características como gênero e idade dos consumidores, e por meio de modelos ARIMA (*Auto Regressive Integrated Moving Average*).

METODOLOGIA

Esta pesquisa é de natureza básica, descritiva, cujo procedimento é um estudo de caso e com o método quantitativo como descrito por Prodanov e Freitas (2013).

Os dados históricos sobre o índice de inadimplência no Serasa foram coletados e organizados em uma série temporal. Para a modelagem matemática, utilizamos técnicas de regressão, considerando o índice de inadimplência como variável resposta e as variáveis gênero e idade dos consumidores como variáveis explicativas. Além disso, empregamos análise do teste *t-student* a fim investigar possíveis diferenças significativas entre grupos.

Essas análises foram realizadas em Python (ATWAN, 2022; HUANG, PETUKHINA, 2022) e o

Ambiente de Desenvolvimento Integrado Jupyter Notebook (JUPYTER, 2023), por possuir diversos repositórios de pacotes disponíveis.

A Serasa (SERASA, 2023) uma empresa brasileira de caráter privado com a missão de coletar informações, realizar pesquisas conduzidas e análises sobre pessoas físicas e jurídica que possuem débitos financeiros. Foi fundada em 1968, como uma iniciativa da Federação Brasileira de Bancos (FEBRABAN), em parceria com a Associação de Bancos do Estado de São Paulo (Assobesp).

O principal objetivo do Serasa é coletar e disponibilizar informações sobre a situação financeira de pessoas físicas e jurídicas, formando um banco de dados que é amplamente utilizado por instituições financeiras, empresas, e até mesmo pelo próprio consumidor. Além de realizar análises detalhadas dos principais indicadores econômicos do país, envolvendo Atividade Econômica e Emprego, Taxas de Juros e de Câmbio, Inadimplência e Crédito geral. Por meio do seu banco de dados, a Serasa oferece informações relevantes para auxiliar na avaliação de riscos financeiros e na tomada de decisões por parte de empresas e instituições financeiras.

Durante a etapa de pré-processamento, os dados coletados foram “limpos” para garantir qualidade e consistência, incluindo imputação de valores ausentes, correção de erros de digitação e padronização dos dados. Então, além de estimar estatísticas descritivas como média, mediana, desvio padrão e percentis, a análise exploratória é realizada usando gráficos de linhas para extrair as características iniciais da série temporal.

Através da decomposição sazonal, identificaram-se e isolaram-se os componentes de tendência, sazonalidade e resíduo na série temporal. Posteriormente, procedeu-se à modelagem dos componentes residuais utilizando-se a metodologia dos modelos ARIMA. A escolha do modelo ARIMA apropriado foi conduzida por meio de análise de autocorrelação, testes de estacionariedade e critérios de seleção de modelos, incluindo o critério de informação de Akaike (AIC). Uma vez ajustado o modelo ARIMA aos dados, realizou-se a validação do mesmo para verificar sua capacidade preditiva.

Tal procedimento foi realizado por meio de uma divisão dos dados em conjuntos de treinamento e teste, ajustando-se o modelo aos dados de treinamento e confrontando as previsões com os dados de teste. Com o modelo ARIMA devidamente validado, efetuaram-se projeções do índice de inadimplência para além do intervalo temporal disponível, fornecendo uma valiosa ferramenta para a estimativa de padrões futuros. Os resultados obtidos revelam a efetividade da abordagem estatística e de modelagem adotada neste estudo.

RESULTADOS E DISCUSSÃO

A discussão dos resultados permitirá que os estudantes compreendam como a modelagem matemática pode ser aplicada na interpretação de dados financeiros complexos, possibilitando a identificação de fatores que influenciam a inadimplência na Serasa. Além disso, os alunos terão a oportunidade de discutir as limitações dos modelos utilizados e a importância de considerar outros fatores que podem afetar a inadimplência.

A estatística descritiva dos dados analisados é mostrada na Tabela 1.

Tabela 1. Resumo da estatística descritiva dos dados analisados.

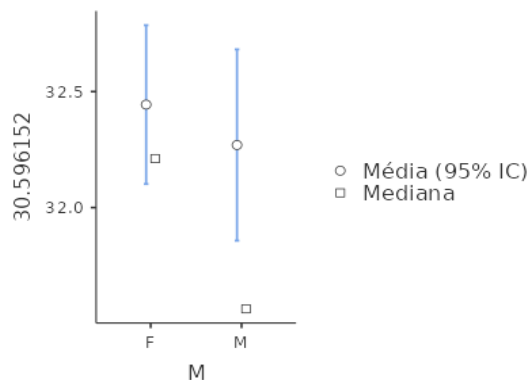
	N	Omissão	Média	Mediana	Desvio-padrão	Mínimo	Máximo
Consumidores Inadimplentes (milhões)	86	0	62.82	62.49	3.204	58.63	71.44
Dívida Média (R\$)	86	0	3982.58	3927.66	214.028	3624.78	4767.20
Masculino	52	34	32.24	31.50	1.506	30.59	35.93
Feminino	52	34	32.44	32.21	1.259	30.78	35.51
Até 25 anos (milhões)	52	34	8.34	8.50	0.483	7.48	9.14
26 - 40 anos	52	34	23.38	23.51	0.810	22.03	25.16
41 - 60 anos	52	34	22.26	21.93	0.885	20.94	24.49
Acima de 60 anos	52	34	10.71	10.55	0.936	9.19	12.65

Fonte: pesquisa própria.

Para investigarmos se há alguma diferença dos grupos masculino e feminino realizamos o teste t para amostras independentes, cujo resultado pode ser visto na Figura 1.

	Estatística	gl	p	Diferença média	Erro-padrão da Diferença
t de Student	0.639	101	0.525	0.174	0.273

Figura 1. Resultado do teste t de Student.



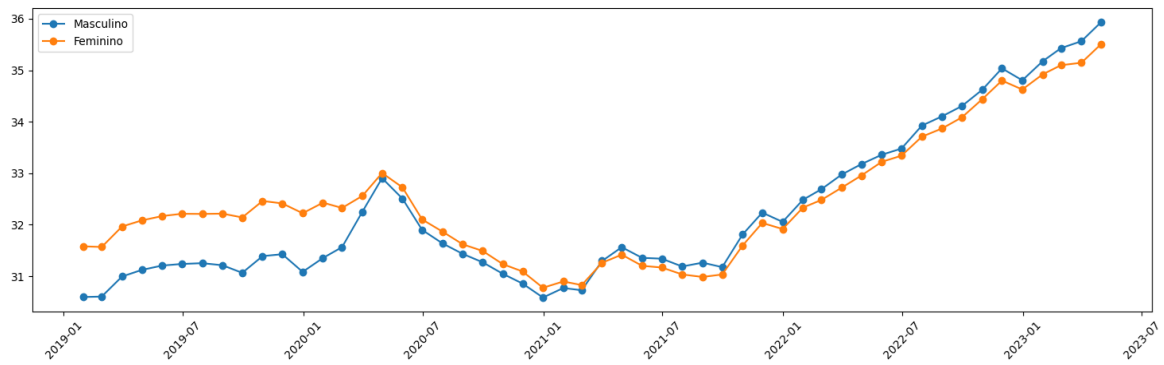
Fonte: própria pesquisa.

Assim, é possível notar que não há diferenças significativas entre os grupos, o que nos leva a concluir que no quesito inadimplência, homens e mulheres se comportam, estatisticamente, iguais. A seguir apresentamos as séries temporais dos inadimplentes e suas principais análises.

Por sexo

Na Figura 2 vemos o número de consumidores inadimplentes (em milhões) no intervalo de 2019 até 2023, em função do sexo.

Figura 2. Série histórica dos inadimplentes em função do sexo (valores em milhões).

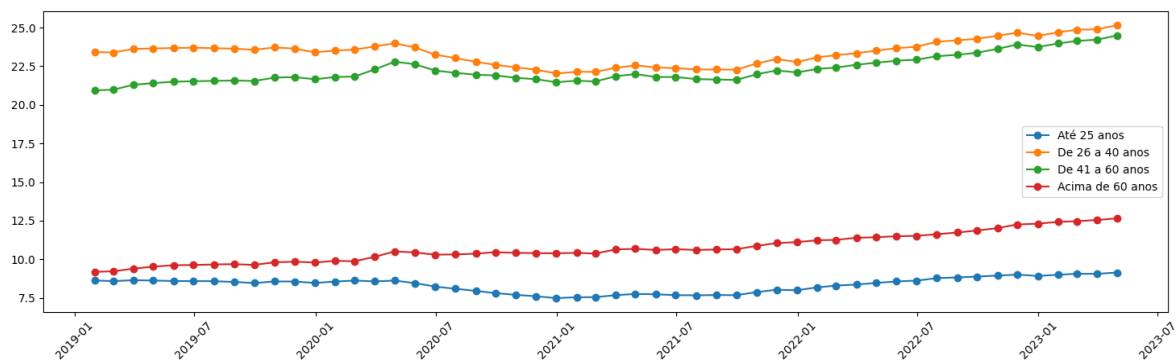


Fonte: própria pesquisa.

Por idade

Na Figura 3 vemos o número de consumidores inadimplentes (em milhões) no intervalo de 2019 até 2023, para algumas faixas etárias.

Figura 3. Série histórica dos inadimplentes em função das idades (valores em milhões).

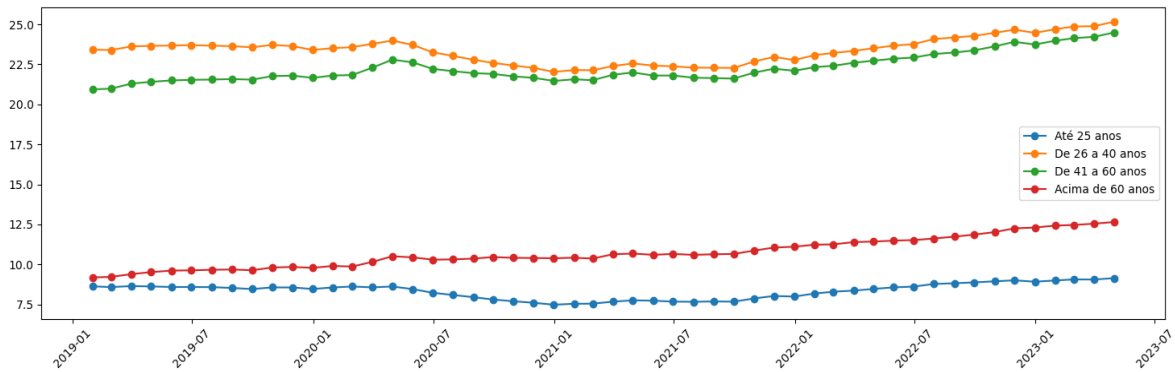


Fonte: própria pesquisa.

É possível notar que as faixas com maior índice de inadimplência é de 26 até 60 anos.

Média móvel

A média móvel é utilizada para suavizar flutuações aleatórias em séries temporais, facilitando a identificação de tendências ou padrões subjacentes. Ela é calculada através da média de um determinado número de observações consecutivas, em função do avanço do tempo. Utilizamos a Média Móvel Simples, que é calculada somando-se um conjunto de valores consecutivos e dividindo o resultado pelo número de valores considerados. Assim, a média móvel para os dados totais são mostrados na Figura 4.

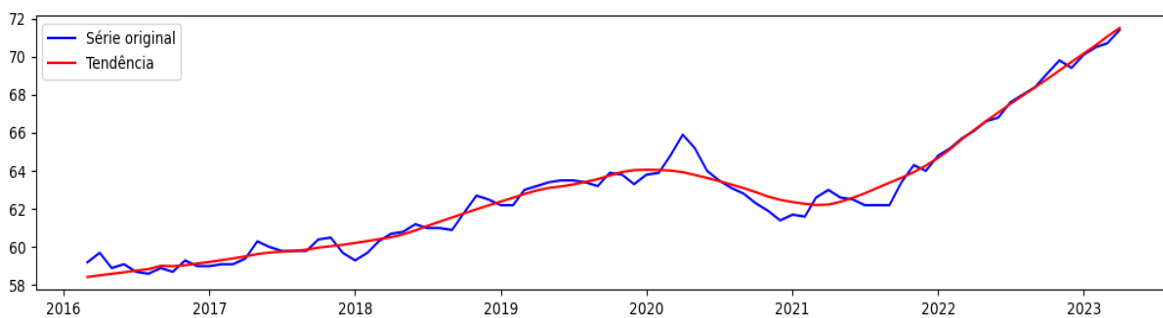
Figura 4. Média móvel dos dados totais.

Fonte: própria pesquisa.

Decomposição

Esta é uma técnica utilizada para analisar e desmembrar uma série temporal em seus componentes principais. Com ela é possível identificar e separar os diferentes elementos presentes da série, como tendência, sazonalidade, variações cíclicas e flutuações aleatórias. A decomposição geralmente envolve três componentes principais.

Tendência: que é a direção dos dados ao longo do tempo, identificando se eles estão aumentando, diminuindo ou mantendo-se relativamente estáveis (Figura 5).

Figura 5. Tendência dos dados totais, com a série original.

Fonte: própria pesquisa.

Sazonalidade: Representa variações regulares e previsíveis que ocorrem em intervalos fixos de tempo, como diariamente, semanalmente, mensalmente ou anualmente (Figura 6).

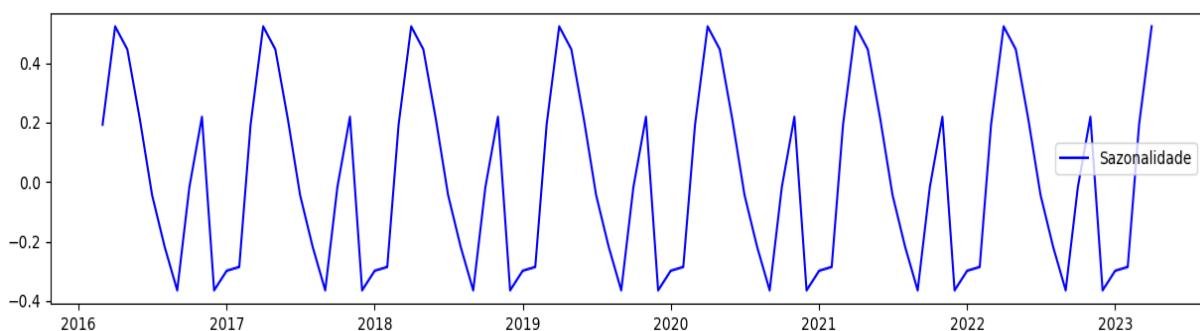
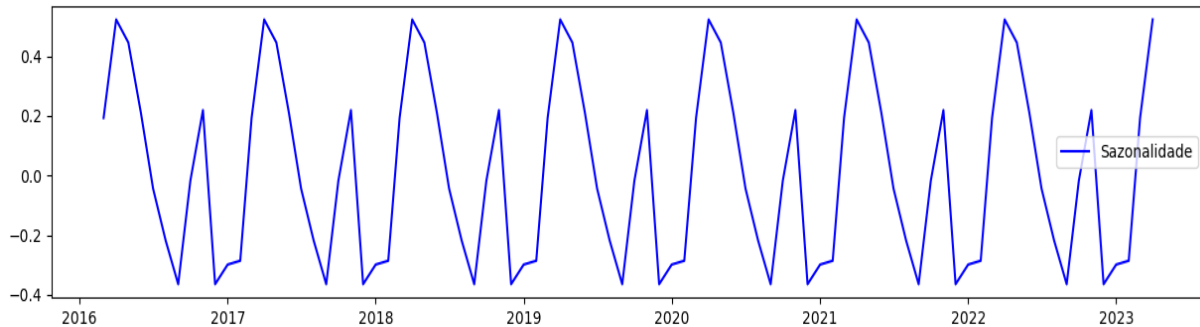


Figura 6. Sazonalidade dos dados totais.

Fonte: própria pesquisa.

Componente residual (ou aleatório): são flutuações não sistemáticas e aleatórias que não podem ser atribuídas à tendência ou à sazonalidade. Essas flutuações podem ser causadas por ruídos, variações aleatórias, fatores imprevisíveis ou erros de medição (Figura 7).

Figura 7. Decomposição residual dos dados totais.

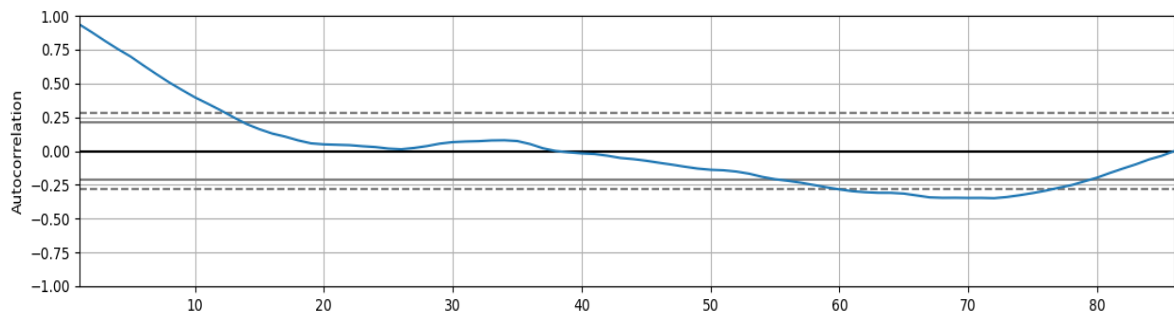


Fonte: própria pesquisa.

Decomposição

Autocorrelação quantifica o grau de correlação entre uma série temporal e suas próprias versões atrasadas, isto é, ela busca uma relação linear entre os valores passados e seus valores atuais. A autocorrelação é utilizada para analisar se existe alguma estrutura de dependência entre os valores ao longo do tempo. O coeficiente de autocorrelação varia de -1 a 1, em que o valor de 1 indica uma correlação perfeita positiva entre os valores, -1 indica uma correlação perfeita negativa e zero indica ausência de correlação.

Figura 8. Autocorrelação dos dados totais.



Fonte: própria pesquisa.

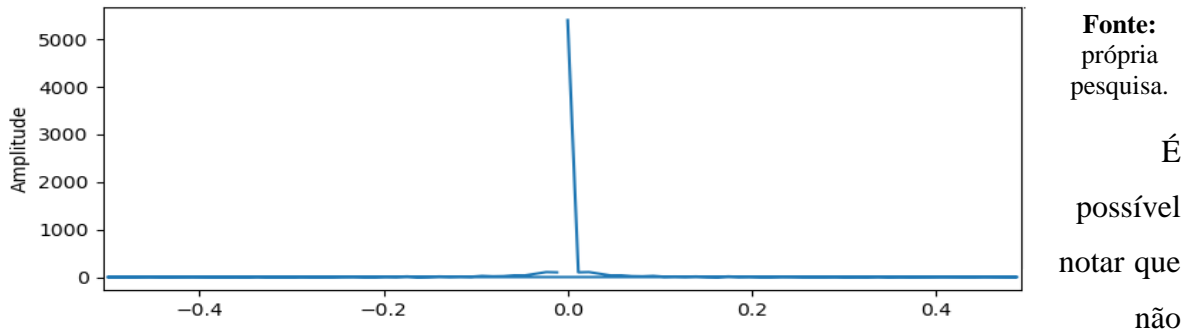
Decomposição espectral

Também conhecida como análise espectral, é uma técnica usada para decompor um sinal ou uma série temporal em suas componentes de frequência. Ela permite identificar as diferentes frequências dominantes nos dados e quantificar a contribuição de cada uma delas.

Uma das principais ferramentas usadas na decomposição espectral é a transformada de Fourier. A

transformada de Fourier converte um sinal do domínio do tempo para o domínio da frequência, representando-o como uma combinação de componentes harmônicos de diferentes frequências. Ela revela os componentes de frequência que compõem o sinal e suas respectivas amplitudes. Utilizamos a chamada transformada rápida de Fourier (FFT) cujos resultados são mostrados na Figura 9.

Figura 9. Decomposição espectral.



existe um valor de frequência dominante na série temporal, a qual indica que os dados não apresentam uma frequência característica regular.

Modelos ARIMA

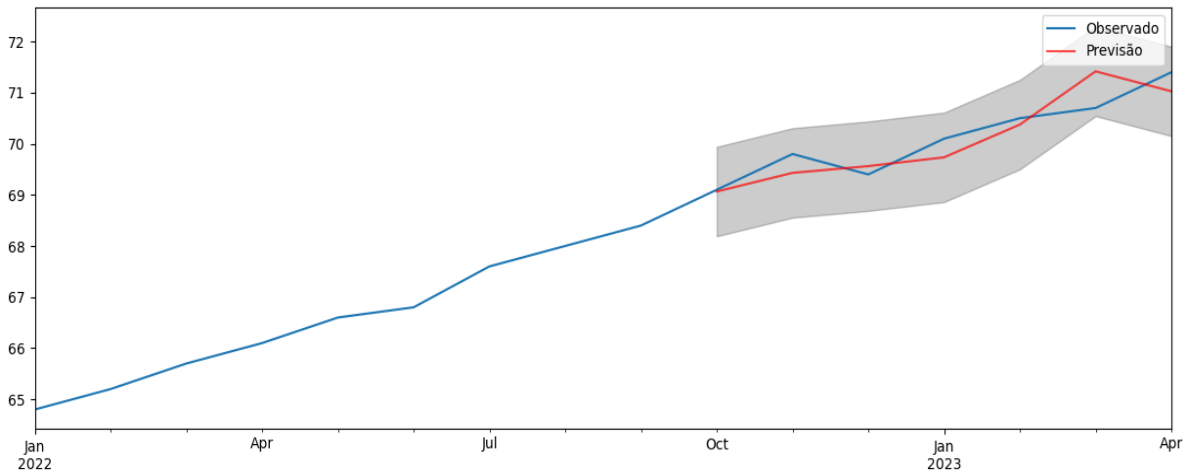
Os modelos ARIMA (*Autoregressive Integrated Moving Average*) são amplamente utilizados na análise e previsão de séries temporais. Esses modelos são compostos por três componentes principais: Um componente autorregressivo (AR), que representa a dependência linear de observações passadas, ou seja, cada valor da série é uma combinação linear de valores anteriores. Portanto, os valores passados de uma série temporal afetam a série atual. O componente integrador (I) refere-se à aplicação de diferenciação a uma série temporal para torná-la estacionária. Esta etapa é necessária porque os modelos ARIMA assumem que a média e a variância da série permanecem constantes ao longo do tempo. O componente de média móvel (MA) descreve a dependência linear de observações anteriores sobre os resíduos do modelo (erros), capturando assim a relação entre os resíduos da série temporal em diferentes pontos no tempo.

O modelo ARIMA é caracterizado por três: p , d e q . O parâmetro p indica a ordem do componente AR, d representa o número de diferenciações integradas aplicadas à série e q representa a ordem do componente MA. A seleção apropriada dos valores de p , d e q é realizada através da análise das funções de autocorrelação e autocorrelação parcial da série temporal. Além desses parâmetros, existem também os parâmetros sazonais P , D e Q , que seguem as mesmas definições dos parâmetros anteriores, porém aplicados aos efeitos sazonais da série.

Depois que um modelo ARIMA é ajustado aos dados históricos, previsões futuras podem ser feitas. Essas previsões são baseadas em uma combinação de componentes AR, I e MA, levando

em considerações estruturas de autocorrelação anteriores. Neste estudo, o modelo ARIMA(1, 1, 1)x(1, 1, 1) [12] foi o mais adequado para os dados analisados, onde [12] representa a periodicidade dos dados. A Figura 10 mostra os resultados da previsão do modelo.

Figura 10. Previsão do modelo ARIMA(1, 1, 1)x(1, 1, 1, 12)[12].



Fonte: própria pesquisa.

É possível notar que, a partir de outubro de 2023, a previsão é bastante satisfatória o que demonstra que os métodos utilizados em nossa análise são adequados para previsões futuras do índice de inadimplentes.

CONSIDERAÇÕES FINAIS

Este exemplo de aplicação destaca a relevância da matemática como uma ferramenta poderosa para a análise de dados financeiros e sociais. A modelagem matemática da inadimplência no Serasa permite aos estudantes explorarem conceitos avançados, como regressão, séries temporais e análise de variância, em um contexto real. A aplicação prática da matemática em problemas do mundo real, como este, contribui para uma formação mais completa e enriquecedora, preparando os futuros professores de matemática para enfrentarem desafios complexos e promovendo uma maior compreensão da importância da matemática nas ciências sociais e financeiras.

REFERÊNCIAS

ATWAN, T. A., Time Series Analysis with Python Cookbook: Practical recipes for exploratory data analysis, data preparation, forecasting, and model evaluation, Packt, Birmingham, 2022.

BRASIL. Ministério da Educação. Base Nacional Comum Curricular. Brasília, 2018.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M., Time Series Analysis: Forecasting and Control, 5a ed., John Wiley & Sons, Inc., Hoboken, New Jersey, 2016.

CRYER, J. D.; CHAN, K.-S., Time Series Analysis: With Applications in R, 2a ed., Springer, New York, 2008.

HUANG, C.; PETUKHINA, A., Applied Time Series Analysis and Forecasting with Python, Springer, Cham, 2022.

JUPYTER, <https://jupyter.org/>, acesso em agos. de 2023.

MORETTIN, P. A., Análise de séries temporais, 2a ed., Blucher, São Paulo, 2006.

PRODANOV, C. C.; FREITAS, E. C. Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico, 2a ed., Universidade Feevale, Novo Hamburgo, 2013.

SERASA, <https://www.serasa.com.br/>, acesso em ago. De 2023.



CERTIFICADO DE ACEITE PARA PUBLICAÇÃO

Data Submissão: **07/08/2023**

Data Aprovação: **14/08/2023**

Certificado Nº: **230813973**

Certificamos que o trabalho intitulado "MODELAGEM MATEMÁTICA DA INADIMPLÊNCIA NO SERASA UMA ANÁLISE ESTATÍSTICA E DE SÉRIE TEMPORAL" de autoria de Júlia Rodrigues Barbosa, Renato Germano foi submetido para avaliação do nosso Conselho Editorial e obteve parecer de **ACEITO para publicação**. O trabalho encontra-se em processo editorial e **será publicado no formato de Capítulo de Livro em 30/09/2023**. A obra será publicada com ISBN, DOI, Conselho Editorial e Declaração de Avaliação.

Reinaldo Cardoso

Editor-Geral **55-518.229/0001-791**

