



**UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE ADMINISTRAÇÃO**

EDICLAUDIO VIEIRA SILVA

**PERFIL DE CLIENTES INADIMPLENTES: ANÁLISE EM UMA EMPRESA DE
PRESTAÇÃO DE SERVIÇOS DE INTERNET DE ARACAJU/SE**

**SÃO CRISTÓVÃO, SE
2024**

EDICLAUDIO VIEIRA SILVA

**PERFIL DE CLIENTES INADIMPLENTES: ANÁLISE EM UMA EMPRESA DE
PRESTAÇÃO DE SERVIÇOS DE INTERNET DE ARACAJU/SE**

Trabalho de Conclusão de Curso (TCC) apresentado ao Departamento de Administração do Centro de Ciências Sociais Aplicadas da Universidade Federal de Sergipe (UFS), como requisito parcial para a obtenção do título de Bacharel em Administração, regulamentadas pela Resolução nº 070/2023/CONEPE.

Orientador(a): Prof. HUANG CHIEN EN

**SÃO CRISTÓVÃO, SE
2024**

EDICLAUDIO VIEIRA SILVA

**PERFIL DE CLIENTES INADIMPLENTES: ANÁLISE EM UMA EMPRESA DE
PRESTAÇÃO DE SERVIÇOS DE INTERNET DE ARACAJU/SE**

**Trabalho de Conclusão de Curso (TCC)
apresentado ao Departamento de
Administração do Centro de Ciências
Sociais Aplicadas da Universidade
Federal de Sergipe (UFS), como
requisito parcial para a obtenção do
título de Bacharel em Administração,
em cumprimento às Normas
regulamentadas pela Resolução no
70/2023/CONEPE.**

Trabalho Defendido e Aprovado em 02 de dezembro de 2024.

BANCA EXAMINADORA

**Prof. Huang Chien En (Orientador)
Universidade Federal de Sergipe (UFS)**

**Prof. Dr. Marcos Eduardo Zambanini (Membro Interno)
Universidade Federal de Sergipe (UFS)**

**Prof. Dr. Maria Elena Leon Olave (Membro Interno)
Universidade Federal de Sergipe (UFS)**

"Seja amorfo, sem forma, como a água. Se você coloca água em uma xícara, ela se torna a xícara. Se coloca em uma garrafa, ela se torna a garrafa. Se a coloca em um bule, ela se torna o bule. A água pode fluir ou pode colidir. Seja como a água, meu amigo."

(Bruce Lee)

AGRADECIMENTOS

Primeiramente, agradeço a Deus por me conceder saúde, força e sabedoria para enfrentar cada etapa desta jornada.

Aos meus pais, Claudio José e Josenilde Vieira, por serem minha base e meu porto seguro. Agradeço profundamente o amor incondicional, a paciência e o apoio em todos os momentos, especialmente nos dias mais difíceis

À minha amada, Maria Clara, por todo o amor, paciência e incentivo ao longo deste percurso. Obrigado por estar ao meu lado, me motivando e acreditando em mim, mesmo nos momentos em que eu próprio duvidei. Sua presença foi essencial para que eu pudesse chegar até aqui.

Aos integrantes da empresa estudada, José Welliton Sá, Paulo Henrique e Jonas Martins, pela autorização e pela valiosa ajuda na confecção e exportação da base de dados. A colaboração de vocês foi fundamental para a realização desta pesquisa e para o sucesso deste trabalho.

Ao meu orientador, Huang Chien En, que, mesmo em licença para doutorado, dispôs-se a me acompanhar e orientar ao longo deste processo. Sua dedicação e suporte foram essenciais para a elaboração deste trabalho e para o meu desenvolvimento como estudante.

RESUMO

O presente Trabalho de Conclusão de Curso tem como objetivo identificar o perfil dos clientes adimplentes e inadimplentes de uma empresa fornecedora de internet em Aracaju/SE. A pesquisa adota uma abordagem quantitativa e descritiva, utilizando dados extraídos do Data Lake (um repositório centralizado onde grandes volumes de dados brutos são armazenados e organizados para posterior análise) da empresa no período de janeiro a dezembro de 2022. Foram analisados fatores como idade, gênero, valor da mensalidade, data de vencimento e localidade dos clientes. A análise foi conduzida utilizando técnicas estatísticas, como o teste t de Student, correlação de Pearson e teste Qui-quadrado, implementados no software R Studio. Esses testes permitiram confirmar a existência de correlações significativas entre variáveis como idade e valor da mensalidade em relação à situação de pagamento dos clientes. Os resultados revelaram que clientes mais jovens e com mensalidades mais elevadas tendem a apresentar maior inadimplência. Por outro lado, clientes mais velhos e com taxas de condomínio mais altas demonstraram uma leve tendência a pagar suas contas com antecedência. A pesquisa contribui para o desenvolvimento de estratégias mais eficazes de gestão financeira e prevenção da inadimplência.

Palavras-chave: Inadimplência, Adimplência, Perfil de Clientes, Análise quantitativa, Empresa de Internet.

ABSTRACT

This Undergraduate Thesis aims to analyze the profile of paying and non-paying customers of an internet service provider in Aracaju/SE. The research adopts a quantitative and descriptive approach, using data extracted from the company's Data Lake (a centralized repository where large volumes of raw data are stored and organized for further analysis) for the period from January to December 2022. Factors such as age, gender, subscription value, payment due date, and customer location were analyzed. The analysis was conducted using statistical techniques, such as Student's t-test, Pearson correlation, and Chi-square test, implemented in the R Studio software. These tests confirmed the existence of significant correlations between variables such as age and subscription value in relation to customers' payment status. The results revealed that younger customers and those with higher subscription values tend to exhibit higher delinquency rates. On the other hand, older customers and those paying higher condominium fees showed a slight tendency to pay their bills in advance. The research contributes to the development of more effective financial management strategies and delinquency prevention.

Keywords: Delinquency Rate, Payment Behavior, Customer Profile, Quantitative Analysis, Internet Service Provider.

LISTA DE FIGURAS

Figura 1– Banco de dados.....	25
Figura 2– Gráfico adimplência e inadimplência x gênero.....	32
Figura 3– Gráfico adimplência e inadimplência por faixa etária.....	33
Figura 4– Gráfico de distribuição por data de vencimento.....	34
Figura 5– Gráfico de distribuição por mensalidade.....	35
Figura 6– Gráfico de distribuição da adimplência e inadimplência por localidade...	35
Figura 7– Fórmula do teste t de student	36
Figura 8– Passos para teste t student do valor da mensalidade no R.....	37
Figura 9– Fórmula do teste Qui-quadrado	42
Figura 10– Teste qui-quadrado para gênero no R.....	42
Figura 11– Fórmula da correlação de Pearson	44
Figura 12– Teste de correlação de Pearson para idade no R.....	44
Figura 13– Teste de correlação de Pearson para taxa de condomínio no R.....	47

LISTA DE QUADROS

Quadro 1– Fatores do perfil dos clientes inadimplentes.....	18
Quadro 2– Processo da descoberta de Conhecimento de Dados.....	20
Quadro 3– Métodos de mineração de dados.....	21
Quadro 4– Variáveis do banco de Dados.....	26
Quadro 5– Distribuição de frequências	29
Quadro 6– Estatística descritiva	30

SUMÁRIO

1 INTRODUÇÃO	12
1.1 PROBLEMA DE PESQUISA	13
1.2 OBJETIVOS	13
1.2.1 Objetivo geral	13
1.2.2 Objetivos específicos	13
1.3 JUSTIFICATIVA	13
2 FUNDAMENTAÇÃO TEÓRICA	15
2.1 INADIMPLÊNCIA E CONSEQUÊNCIAS	15
2.2 INADIMPLÊNCIA DAS FAMÍLIAS BRASILEIRAS	16
2.3 PERFIL DOS CLIENTES INADIMPLENTES	18
2.4 KDD - KNOWLEDGE DISCOVERY IN DATABASES	19
2.5 MINERAÇÃO DE DADOS	21
2.6 OUTROS ESTUDOS SOBRE INADIMPLÊNCIA	22
2.6.1 Estudo de Baldi e Gianetti (2018)	22
2.6.2 Estudo de Higuchi (2023)	23
2.6.3 Estudo de Ferreira (2020)	23
2.6.4 Estudo de Dynan, Skinner e Zeldes (2000)	23
2.6.5 Contribuições dos Estudos	24
3 METODOLOGIA	24
3.1 CARACTERIZAÇÃO DA PESQUISA	24
3.3 VARIÁVEIS UTILIZADAS	26
3.4 ANÁLISE DOS DADOS E HIPÓTESES	27
3.4.1 Metodologia para Teste de Hipóteses	28
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS	29
4.1 ESTATÍSTICA DESCRITIVA	29
4.1.1 Análise da estatística descritiva	31
4.1.2 Tópicos Relevantes da estatística descritiva e distribuição de frequência.....	31
4.2 HIPÓTESES	36
4.2.1 Situação do cliente por valor da mensalidade	36
4.2.1.1 Interpretação dos resultados fornecidos	38
4.2.1.2 Decisão	40
4.2.1.3 Conclusão	41

4.2.2 Situação do cliente por gênero	41
4.2.2.1 Interpretação dos resultados fornecidos	42
4.2.2.2 Decisão	43
4.2.3 Situação do cliente por idade	43
4.2.3.1 Interpretação dos resultados fornecidos	45
4.3 SITUAÇÃO DO CLIENTE POR TAXA DE CONDOMÍNIO	46
4.3.1 Coleta dos dados	46
4.3.2 Formulação do teste	47
4.3.3 Interpretação dos resultados fornecidos	48
4.3.4 Decisão	50
4.4 COMPARAÇÃO DOS RESULTADOS COM OUTROS ESTUDOS	50
4.4.1 Comparação com o Estudo de Baldi e Gianetti (2018)	50
4.4.2 Comparação com o Estudo de Higuchi (2023)	50
4.4.3 Comparação com o Estudo de Ferreira (2020)	51
4.4.4 Comparação com o Estudo de Dynan, Skinner e Zeldes (2000)	51
4.4.5 Síntese Comparativa	52
5 CONCLUSÃO	53
5.1 PRINCIPAIS RESULTADOS	53
5.2 LIMITAÇÕES E SUGESTÕES PARA FUTURAS PESQUISAS	54
5.3 CONSIDERAÇÕES FINAIS	55
REFERÊNCIAS	57

1 INTRODUÇÃO

Um dos principais desafios financeiros enfrentados pelas empresas é o controle do fluxo de caixa, sendo a inadimplência uma dificuldade frequentemente destacada.

Sandroni (1999) descreve a inadimplência como a não observância das disposições contratuais dentro do prazo estipulado. Além disso, a parte que não cumpre suas obrigações contratuais fica sujeita a pagar juros de mora, multa contratual ou outros encargos.

Na troca do produto ou da prestação do serviço a empresa espera que o cliente cumpra com as obrigações contratadas. Contudo, algumas vezes a relação entre a organização e o cliente não é contemplada de maneira acordada entre as partes, o descumprimento de deveres é um assunto que acontece regularmente na sociedade brasileira. (Pinto; Coronel, 2013).

A falta ou atraso nos recebimentos pode causar impacto devastador nas finanças de uma empresa e prejudicar sua reputação no mercado. A inadimplência é um dos principais fatores para o desequilíbrio financeiro segundo a Administra Brasil (2018) é crucial contar com a quitação das dívidas pendentes por parte de seus clientes. Caso ocorra algum imprevisto que prejudique essa previsão, sua capacidade financeira será afetada, comprometendo o fluxo de caixa.

O Brasil possui um elevado número de pessoas endividadas, ou seja, com compromissos financeiros a pagar. Entre essas, muitas acabam inadimplentes, o que significa que deixam de honrar suas dívidas no prazo estipulado. Essa inadimplência contribui para a instabilidade econômica do país, afetando tanto as empresas quanto o governo. Segundo Pesquisa Confederação Nacional do Comércio de Bens, Serviços e Turismo (2023), as dívidas alcançam 78% das famílias brasileiras sendo 29% com dívidas em atraso e 12% sem condições de quitar suas dívidas, com crescimento anual de 7 pontos percentuais, demonstrando aceleração nos pós pandemia. Os principais atingidos foram mulheres jovens, consumidor de baixa renda, com mais idade e menos escolaridade.

“A falta de pagamento causa muito transtorno às empresas e por isso, é preciso conhecê-la e saber quais os fatores que ocasionam a inadimplência, criando meios para se prevenir dos problemas que ela pode vir a causar.” (Andrade, et al.,2008 p.47).

A partir disso, a empresa enfrenta desafios na administração de seus recursos financeiros e na entrega de produtos e serviços quando lida com clientes que são considerados inadimplentes. Portanto, é importante compreender o perfil dos clientes em situação de inadimplência e as razões que os levaram a essa condição. Essa análise permitirá que a empresa adote medidas preventivas ou mitigadoras para minimizar as repercussões negativas causadas pela inadimplência. (Pinto; Coronel, 2013).

1.1 PROBLEMA DE PESQUISA

Qual o perfil dos clientes adimplentes e inadimplentes em uma empresa fornecedora de internet de Aracaju/SE?

1.2 OBJETIVOS

1.2.1 Objetivo geral

O objetivo geral deste trabalho é identificar o perfil dos clientes adimplentes e inadimplentes de uma empresa fornecedora de internet em Aracaju/SE.

1.2.2 Objetivos específicos

- Verificar possíveis fatores, que possam estar relacionados a adimplência.
- Verificar possíveis fatores, que possam estar relacionados a inadimplência.
- Correlacionar dados dos clientes adimplentes e inadimplentes.

1.3 JUSTIFICATIVA

A inadimplência é um desafio crítico enfrentado por empresas de diversos setores, e o setor de telecomunicações não é exceção. No contexto de uma empresa de telecomunicações, do ponto de vista de Gomes (2013), a inadimplência de clientes se apresenta como um fator preocupante e desafiador para as empresas do ramo. Com efeito, essa problemática impõe à área de crédito e cobrança a responsabilidade de empregar estratégias eficazes, com o propósito de reduzir o número de inadimplentes de forma contínua. Somente assim, as organizações podem garantir

sua competitividade em um mercado acirrado e, ao mesmo tempo, sustentar sua lucratividade em um cenário econômico cada vez mais complexo e competitivo.

A relevância desta pesquisa reside na sua contribuição para uma melhor compreensão dos padrões e que levam os clientes a tornarem-se inadimplentes em serviços de internet. Ao entender o perfil dos clientes inadimplentes e comparar com o perfil dos adimplentes, a empresa pode desenvolver estratégias mais eficazes para prevenir a inadimplência, melhorar sua gestão financeira e fortalecer o relacionamento com os consumidores.

Embora existam diversos estudos sobre inadimplência no setor de telecomunicações, este trabalho foca em compreender os desafios e estratégias específicas de provedores de internet regionais, com ênfase no contexto de Aracaju-SE. Os resultados obtidos podem oferecer contribuições valiosas para o aprimoramento das práticas de gestão financeira nesse setor e servir como referência para profissionais e pesquisadores interessados em soluções aplicadas ao cenário local

A importância dessa pesquisa também está relacionada ao seu potencial impacto econômico e social. Ao desenvolver estratégias mais efetivas de prevenção da inadimplência, a empresa estudada pode reduzir suas perdas financeiras, melhorar a qualidade do serviço prestado e aumentar a satisfação dos clientes. Por consequência, isso pode refletir positivamente no mercado de telecomunicações, estimulando a concorrência saudável e beneficiando os consumidores em geral.

Segundo a Agência Nacional de Telecomunicações (2023), o Brasil possui mais de 200 mil empresas prestadoras de serviços de telecomunicações, sendo 2.459 em Sergipe. Este estudo focou em uma dessas empresas devido à viabilidade de acesso aos dados necessários para a análise, bem como à relevância do seu contexto operacional no mercado local de Aracaju/SE. Assim, justifica-se a escolha desta empresa específica pelo seu alinhamento com os objetivos da pesquisa e pela oportunidade de colaborar com estratégias voltadas para a melhoria da gestão financeira e a redução da inadimplência no setor.

2. FUNDAMENTAÇÃO TEÓRICA

A inadimplência, é um desafio significativo enfrentado por empresas de diversos setores. A presente seção do trabalho busca aprofundar a compreensão desse fenômeno complexo, abordando suas causas e consequências, estratégias de prevenção e gerenciamento, bem como seu impacto específico no contexto do setor de telecomunicações.

2.1 INADIMPLÊNCIA E CONSEQUÊNCIAS

A inadimplência é um termo que se refere à falta de cumprimento de obrigações financeiras acordadas em um contrato ou acordo. Em outras palavras, quando uma pessoa, empresa ou entidade deixa de fazer um pagamento ou cumprir uma obrigação financeira dentro do prazo estipulado, ela é considerada inadimplente. No contexto empresarial, a inadimplência pode envolver o não pagamento de contas, dívidas, empréstimos ou parcelas de financiamento conforme os termos acordados. Isso pode resultar em consequências financeiras, como a aplicação de juros de mora, multas contratuais e a possibilidade de acionar medidas legais para recuperação dos valores devidos, (Barth, 2004; Aziz, 2023; Matias; Daubermann e Rici, 2009).

Por outro lado, adimplência é o termo utilizado para descrever o estado em que uma pessoa, empresa ou entidade cumpre regularmente com suas obrigações financeiras de acordo com os prazos estipulados em contratos ou acordos. Quando uma parte é considerada adimplente, isso indica que ela está em dia com seus pagamentos, dívidas ou outras responsabilidades financeiras, o que geralmente é visto como um sinal de confiabilidade e boa gestão financeira. A adimplência pode resultar em benefícios, como a manutenção de uma boa reputação creditícia, condições mais favoráveis em negociações financeiras e a redução do risco percebido pelos credores. A manutenção da adimplência é fundamental para estabelecer relações financeiras saudáveis e evitar as implicações negativas associadas à inadimplência (Barth, 2004; Claudino, 2009; Aziz, 2023; SPC, 2016).

Como consequência a inadimplência traz diversos prejuízos para as organizações credoras. Segundo Tavares (1996), a inadimplência coloca a instituição em uma posição desvantajosa, levando a tesouraria a explorar alternativas para garantir um fluxo contínuo de recursos. Essa situação cria um impacto negativo nas finanças da empresa devido a diversos fatores, como a presença constante da

inadimplência, a incerteza associada a ela e a sua durabilidade ao longo do tempo. Esses elementos, combinados, trazem desafios adicionais para a administração financeira da organização.

Tavares (1996) ainda complementa que em um cenário caracterizado por taxas inflacionárias, a inadimplência assume um peso ainda maior. Ela resulta em uma perda substancial no valor do capital não recebido pela empresa credora, pois a inflação diminui o poder de compra do dinheiro ao longo do tempo. A incapacidade de receber o capital devido no prazo acordado não apenas impacta negativamente a liquidez imediata da empresa, mas também cria um efeito cascata que afeta a capacidade da organização de cumprir compromissos financeiros futuros.

2.2 INADIMPLÊNCIA DAS FAMÍLIAS BRASILEIRAS.

O Brasil possui um grande número de endividados o que leva posteriormente a inadimplência na qual é influenciada por fatores econômicos, psicológicos e sociológicos (Barth, 2004).

A expressão "inadimplência" ganhou maior notoriedade no Brasil nos anos 1990, em decorrência de importantes transformações na economia do país. As significativas mudanças econômicas que se seguiram ao conjunto de medidas implementadas pelo então Presidente Fernando Collor de Melo, com o objetivo de combater a inflação e forçar a redução de preços por parte dos empresários, afetaram profundamente o poder de compra dos consumidores. Isso ocorreu devido à drástica redução na circulação de moeda e ao congelamento das contas correntes, poupanças e investimentos (Sebben, 2020 *apud* Reis, 2020). Estudiosos buscam entender e explicar os fatores, abrangendo tanto aspectos macroeconômicos quanto microeconômicos, que contribuem para a inadimplência na sociedade brasileira. Seguem algumas dessas explicações:

Para Vieira *et al.* (2014) A diminuição da taxa Selic sinaliza períodos em que as condições para acessar crédito se tornam mais favoráveis, criando assim uma maior disponibilidade de crédito que, por sua vez, tem o potencial de impulsionar o aumento do endividamento, contribuindo para o surgimento de casos de inadimplência. Dessa forma, é possível estabelecer uma correlação inversa entre a taxa Selic e o nível de inadimplência observado nas famílias brasileiras,

Já os autores Silva et al. (2020), Caetano et al (2016), Baumann (2008) e Carvalho (2019), defendem que um grande fator para o endividamento é o consumismo causado por influências do marketing e propaganda, pressões sociais, fetichização da mercadoria, facilidade de crédito, busca por satisfação imediata e a necessidade de autoafirmação econômica. Esses fatores levam as pessoas a comprarem em excesso, frequentemente além de suas possibilidades financeiras, resultando em endividamento e outras consequências negativas.

A inadimplência surge quando há falhas no cumprimento de compromissos financeiros, seja por falta de planejamento, mudanças inesperadas na renda ou outros fatores (Tavares, 1996). Em alguns casos, o endividamento excessivo pode levar à inadimplência, especialmente quando os recursos financeiros disponíveis não são suficientes para cobrir todas as dívidas acumuladas (SPC BRASIL, 2016). Da mesma forma, a inadimplência pode agravar o endividamento, devido à acumulação de juros e encargos sobre pagamentos atrasados (Bath, 2004).

Os autores Antoniazi (2019), Zerrenner (2007) e Ruberto et al (2013), relatam que o desemprego desempenha um papel central no fenômeno do endividamento no Brasil. Quando um indivíduo perde seu emprego, enfrenta uma considerável redução em sua fonte de renda. Isso pode acarretar dificuldades financeiras, uma vez que as despesas diárias persistem, mesmo que a renda tenha sido reduzida ou eliminada. Para enfrentar essa diminuição na renda, muitas pessoas se veem compelidas a buscar crédito, seja por meio de empréstimos bancários, cartões de crédito ou financiamentos, a fim de atender às necessidades básicas, como habitação, alimentação e educação.

A imperatividade de recorrer ao crédito em razão do desemprego frequentemente resulta em um aumento no nível de endividamento, uma vez que as pessoas acumulam dívidas para manter seu padrão de vida anterior ou para enfrentar situações financeiras imprevistas. Adicionalmente, as elevadas taxas de juros associadas a muitos desses instrumentos de crédito contribuem para o aumento do montante total das dívidas ao longo do tempo (Amaral, 2015)

De acordo com os autores Ferreira (2017) e Claudino (2009), A falta de educação financeira também está intimamente ligada à inadimplência pois as pessoas que não possuem conhecimentos financeiros adequados podem tomar decisões financeiras ruins. Isso inclui contrair dívidas sem entender completamente os custos

envolvidos, gastar mais do que podem e não ter um planejamento financeiro adequado. A educação financeira desempenha um papel crucial na prevenção da inadimplência, capacitando as pessoas a tomar decisões informadas, gerenciar suas finanças com responsabilidade e evitar armadilhas financeiras que levam ao endividamento excessivo. Portanto, promover a educação financeira é fundamental para reduzir a inadimplência e ajudar as pessoas a manterem uma situação financeira saudável.

2.3 PERFIL DOS CLIENTES INADIMPLENTES

Conforme enfatizado por Bail (2014), além de manter um relacionamento sólido com seus clientes, a empresa deve regularmente examinar o comportamento deles em relação às pendências financeiras com a organização, visando aprimorar a tomada de decisões com maior eficácia e eficiência. Por Meio dessa análise de inadimplência, é possível identificar elementos no perfil dos clientes que possuem dívidas em aberto, resultando na criação de um conjunto de informações que servirá como base para definir estratégias de ação para o departamento de cobrança. O Quadro 1 traz elementos para serem analisados e descobrir o perfil dos clientes:

Quadro 1 – Fatores do Perfil dos Clientes Inadimplentes

NUMERO	ELEMENTO	DESCRIÇÃO	AUTOR
1	FAIXA ETÁRIA	Separar os grupos por faixa etária e identificar padrões de inadimplência	Spc Brasil (2016), Serasa (2022), Viana, Silva e Santos (2019).
2	GENÊRO	Analisar se o gênero tem influência na inadimplência	Spc Brasil (2016), Serasa (2022), Ganske (2020).
3	DIAS DE ATRASO	Identificar padrões e média de quantos dias os clientes atrasam seu pagamento	Ganske (2020), Viana, Silva e Santos (2019), Rocha (2021).
4	LOCALIDADE	Verificar a influência da localização geográfica dos clientes, e quais locais possuem cliente mais inadimplentes	Albuquerque, Medina e Silva (2017); Tochetto (2018).
5	VALOR MENSALIDADE	Verificar quais mensalidades possuem mais inadimplência	Viana, Silva e Santos (2019).

Fonte: Elaborado pelo autor, 2023

Após a definição dos elementos a serem analisados, torna-se essencial realizar uma coleta de dados abrangente e minuciosa, com o objetivo de obter informações detalhadas e abrangentes sobre o perfil dos clientes inadimplentes. Essa coleta de dados desempenha um papel fundamental na compreensão profunda das razões subjacentes à inadimplência, permitindo, assim, uma abordagem mais informada e eficaz na gestão desses clientes e na implementação de estratégias mais bem-sucedidas, (Guterres, 2009).

Essas estratégias, quando bem implementadas, não apenas auxiliam na gestão eficaz dos clientes inadimplentes, mas também fornecem subsídios para a formulação de medidas preventivas e corretivas (Guterres, 2009). A seguir, serão discutidas as metodologias utilizadas para analisar o perfil dos clientes e os resultados obtidos, destacando os fatores determinantes para a inadimplência.

2.4 KDD - KNOWLEDGE DISCOVERY IN DATABASES

No mundo digitalizado e conectado à internet, a quantidade crescente de dados é enorme e requer soluções eficazes de armazenamento e processamento. A análise manual é inviável devido à escala e complexidade dos dados, tornando essencial a automação com métodos computacionais para extrair informações valiosas. A computação permite processar e analisar dados de forma eficiente, identificando padrões e tendências ocultas. Portanto, a integração de métodos computacionais é fundamental para lidar com o crescente volume e complexidade de dados, transformando-os em insights que impulsionam decisões e estratégias bem-sucedidas (Fayyad et al., 1996; Han e Kamber, 2012; Witten, Eibe e Hall, 2011).

Dentre esses métodos, destaca-se o KDD (Knowledge Discovery in Databases) – traduzido como Descoberta de Conhecimento em Bancos de Dados – que é um processo complexo para identificar padrões em dados que sejam compreensíveis e potencialmente valiosos. Conforme Han, Pei e Kamber (2022), as etapas do KDD incluem desde a limpeza e integração dos dados até a seleção de padrões e a interpretação dos resultados, permitindo que grandes volumes de dados sejam transformados em conhecimento útil para a tomada de decisão.

Os avanços recentes, como discutido por Silva et al. (2023), incorporaram ferramentas baseadas em aprendizado profundo e inteligência artificial ao KDD, promovendo análises mais robustas e preditivas. Essas inovações são especialmente

úteis em aplicações modernas, como sistemas de recomendação, previsões financeiras e diagnósticos médicos baseados em dados. Esses avanços consolidam o KDD como um processo contínuo de adaptação e evolução, alinhado às necessidades crescentes de análise de dados em escala global. A seguir, o Quadro 2 detalha as etapas do processo de KDD, conforme descrito por Fayyad et al. (1996), complementadas pelas aplicações e metodologias modernas discutidas por Han, Pei e Kamber (2022) e Silva et al. (2023).

Quadro 2 – Processo da Descoberta de Conhecimento de Dados

NUMERO	PROCESSOS	DESCRIÇÃO
1	Preparação dos Dados	Envolve adquirir uma compreensão aprofundada das informações pertinentes para solucionar o problema em questão, bem como determinar o objetivo do processo KDD sob a perspectiva do utilizador final.
2	Estabelecer o Conjunto de Dados	Consiste a seleção de um conjunto de dados particular ou a concentração em um subconjunto de variáveis ou amostras de dados que serão priorizados durante o processo de KDD.
3	Limpeza e Pré-processamento de Dados	Engloba a eliminação de possíveis interferências ou "ruídos", a obtenção das informações necessárias para modelagem ou tratamento desse ruído, decisões sobre como lidar com campos de dados ausentes e a consideração de informações de sequenciamento temporal e mudanças conhecidas
4	Redução dos Dados	Envolve a aplicação de técnicas para reduzir a dimensionalidade dos dados, o que significa que o número de variáveis em consideração pode ser reduzido dependendo do método utilizado
5	Mineração de Dados	Requer a determinação da tarefa de Mineração de Dados necessária para responder à pergunta de negócio, ou seja, a escolha entre técnicas de regressão, classificação ou agrupamento. Além disso, é crucial selecionar o algoritmo mais adequado dentro da tarefa escolhida para buscar padrões no conjunto de dados.
6	Interpretação dos Resultados	Esta fase engloba a visualização dos padrões extraídos e, conseqüentemente, a interpretação adequada desses resultados. Isso leva à obtenção do conhecimento derivado do banco de dados, o qual será empregado para apoiar futuras tomadas de decisão.

Fonte: Fayyad, Piatetsky-Shapiro e Smyth (1996).

Os conceitos apresentados no quadro destacam as principais etapas do processo de KDD, que incluem desde a preparação e integração de dados até a interpretação dos resultados. Dentre essas etapas, a mineração de dados se sobressai como uma das mais importantes, pois é nessa fase que os padrões e relações ocultas nos dados são efetivamente descobertos (Han e Kamber, 2012). Devido à sua relevância prática e teórica, a mineração de dados é tratada como

um tópico independente neste trabalho, permitindo uma discussão mais aprofundada sobre seus métodos, aplicações e implicações no contexto da descoberta de conhecimento.

2.5 MINERAÇÃO DE DADOS

Ao longo do processo de Descoberta de Conhecimento em Bancos de Dados (KDD), uma das fases críticas é a Mineração de Dados, na qual os algoritmos são aplicados aos dados coletados. Essa aplicação requer uma avaliação cuidadosa e um planejamento sólido, considerando os tipos de dados disponíveis e a questão inicial que motivou o estudo (Han, Pei e Kamber, 2022)

Os autores Han e Kamber (2006) descrevem os tipos de padrões que podem ser identificados nos dados, como associações, sequências temporais, classificações e agrupamentos. Silva et al. (2023) destacam que as técnicas modernas de mineração de dados, como aprendizado profundo e algoritmos de inteligência artificial, têm ampliado a capacidade de identificar padrões complexos e relevantes, especialmente em grandes volumes de dados. Durante a fase de mineração, essas técnicas exploram associações, classificações e agrupamentos, permitindo análises mais robustas e preditivas. No Quadro 3, estão descritos os principais métodos de mineração de dados, suas aplicações e características.

Quadro 3- Métodos de mineração de dados

NUMERO	MÉTODOS	DESCRIÇÃO
1	ASSOCIAÇÕES	A mineração de associações envolve a descoberta de relações interessantes entre itens em um conjunto de dados. Ela é frequentemente usada para encontrar regras do tipo "se X, então Y", onde X e Y são conjuntos de itens. Essas regras descrevem a frequência com que certos itens coocorrem em um conjunto de dados.
2	CLASSIFICAÇÕES	A classificação é uma técnica de mineração de dados usada para prever a classe ou categoria de um objeto com base em suas características. É uma técnica supervisionada, o que significa que ela requer um conjunto de dados de treinamento com exemplos rotulados.
3	REGRESSÃO	A regressão é outra técnica supervisionada usada para prever um valor numérico (variável de saída) com base em um ou mais atributos (variáveis de entrada). Ela modela a relação entre as variáveis de entrada e a saída para fazer previsões.
4	AGRUPAMENTO	Envolve a identificação de grupos (clusters) de objetos similares em um conjunto de dados, onde os objetos dentro do mesmo grupo são mais similares entre si do que com os objetos de outros grupos.

Fonte: Han e Kamber (2006); Silva et al. (2023).

Cada uma dessas técnicas desempenha um papel fundamental na análise de dados e na extração de informações valiosas para tomada de decisões em diversas áreas, como negócios, ciência, medicina e muito mais. A escolha da técnica a ser utilizada depende da natureza dos dados e dos objetivos da análise, Han e Kamber (2006).

2.6 OUTROS ESTUDOS SOBRE INADIMPLÊNCIA

A inadimplência é um fenômeno amplamente estudado em diferentes contextos e setores, sendo influenciada por diversas variáveis socioeconômicas e comportamentais. A seguir, são analisados estudos que abordam fatores determinantes da inadimplência em diferentes contextos, fornecendo uma base teórica relevante para a identificação de clientes adimplentes e inadimplentes em uma empresa de internet em Aracaju/SE.

2.6.1 Estudo de Baldi e Gianetti (2018)

Baldi e Gianetti (2018) investigaram os fatores determinantes da inadimplência em uma cooperativa de crédito, com foco em variáveis como gênero, nível de instrução e localização. O estudo concluiu que as mulheres apresentaram uma menor probabilidade de inadimplência em comparação aos homens. Além disso, indivíduos com maior nível de escolaridade também mostraram uma tendência a serem adimplentes, enquanto a renda não se apresentou como um fator significativo para a probabilidade de inadimplência.

A conclusão de Baldi e Gianetti (2018) destaca a importância de variáveis socioeconômicas, como a educação, na análise de comportamento financeiro. Embora a renda não tenha demonstrado correlação significativa, a educação parece desempenhar um papel crucial no cumprimento das obrigações financeiras, indicando que indivíduos mais instruídos tendem a adotar uma postura mais responsável em relação a seus pagamentos.

2.6.2 Estudo de Higuchi (2023)

Higuchi (2023) conduziu um estudo focado em uma *fintech* de meios de pagamento, explorando correlações entre características dos clientes e inadimplência. O estudo identificou que a maior parte dos inadimplentes era composta por clientes com idade média de 43 anos, localizados majoritariamente na região Sudeste do Brasil. O método de pagamento também foi relevante, com maior inadimplência associada ao uso de cartões de crédito em comparação com outros meios de pagamento.

Este estudo sugere que variáveis demográficas, como a idade e a localização geográfica, podem influenciar significativamente o comportamento de inadimplência.

2.6.3 Estudo de Ferreira (2020)

Ferreira (2020) analisou as causas da inadimplência entre profissionais de enfermagem no Conselho Regional do Amazonas, dividindo as variáveis em fatores microeconômicos e macroeconômicos. O estudo destacou que os profissionais com maior renda tendiam a apresentar um maior comprometimento salarial, o que os tornava mais propensos à inadimplência. Além disso, a carência de conhecimentos em educação financeira foi identificada como um fator crítico para o aumento da inadimplência.

O estudo de Ferreira (2020) reforça a relevância da gestão financeira pessoal na prevenção da inadimplência, evidenciando que o comprometimento financeiro exacerbado, mesmo entre indivíduos de maior renda, pode levar a dificuldades no pagamento de dívidas. O conhecimento inadequado de práticas de educação financeira surge como um fator determinante para a inadimplência.

2.6.4 Estudo de Dynan, Skinner e Zeldes (2000)

Dynan, Skinner e Zeldes (2000) exploraram de que forma as despesas fixas, como moradia, alimentação e outros custos recorrentes, afetam a capacidade de poupança e o comportamento financeiro das famílias. O estudo identificou que despesas fixas consomem uma parte significativa da renda disponível, limitando a flexibilidade dos consumidores para lidar com imprevistos e cumprir outras obrigações financeiras.

Esse estudo é especialmente relevante na análise de como despesas fixas podem impactar o comportamento de inadimplência. Ao consumir uma parte expressiva da renda, essas despesas podem reduzir a capacidade de os consumidores honrarem outros compromissos financeiros, sugerindo que as obrigações financeiras fixas são determinantes importantes para a inadimplência.

2.6.5 Contribuições dos Estudos

Os estudos revisados fornecem uma série de elementos importantes para a compreensão do fenômeno da inadimplência. Baldi e Gianetti (2018) mostram que variáveis como gênero e educação influenciam o comportamento de pagamento, enquanto Higuchi (2023) destaca o impacto da idade e do método de pagamento sobre a inadimplência. Ferreira (2020) traz à tona a importância da educação financeira e do gerenciamento adequado da renda, especialmente para evitar o comprometimento salarial excessivo. Por fim, Dynan, Skinner e Zeldes (2000) ressaltam o papel das despesas fixas como fator limitador da flexibilidade financeira, impactando diretamente a capacidade de adimplência dos indivíduos.

Esses estudos ajudam a fundamentar teoricamente a análise da inadimplência em diferentes contextos, fornecendo subsídios para o entendimento dos fatores que influenciam o comportamento financeiro dos clientes em setores diversos, incluindo o de serviços de telecomunicações.

3. METODOLOGIA

Nesta seção, são descritos os métodos e procedimentos que foram utilizados para analisar as relações entre o perfil dos clientes adimplentes e inadimplentes na empresa estudada.

3.1 CARACTERIZAÇÃO DA PESQUISA

A pesquisa adota a natureza quantitativa e descritiva. Segundo Gil (2002), uma abordagem quantitativa é uma metodologia de pesquisa que se concentra na coleta e análise de dados numéricos, com o objetivo de medir, quantificar e identificar relações estatísticas entre variáveis.

O autor ainda cita que a análise estatística desempenha um papel fundamental na pesquisa quantitativa, permitindo que os pesquisadores identifiquem tendências, relações, padrões e diferenças significativas nos dados coletados. Isso é crucial para a validação das hipóteses formuladas na pesquisa quantitativa.

As pesquisas descritivas têm como principal objetivo descrever as características de uma população ou fenômeno específico. Elas também podem ser usadas para identificar possíveis relações entre variáveis. Essas pesquisas são muito comuns e muitas das pesquisas realizadas no âmbito profissional se enquadram nessa categoria. Dentro das pesquisas descritivas, destacam-se aquelas que investigam as características de grupos, como idade, gênero, origem, nível de educação, saúde física e mental, entre outros (Gil 2010).

3.2 A BASE DE DADOS

Os dados utilizados como base para as análises do estudo foram extraídos do Data Lake da empresa por meio do sistema IXC Soft, um software de gestão empresarial voltado para provedores de internet e empresas de telecomunicações, no período de janeiro de 2022 a dezembro de 2022. Os registros possuem 36.412 linhas e as seguintes colunas: Nome do cliente, CPF, data da emissão, data do vencimento, data do pagamento, dias de atraso, situação do pagamento, valor recebido, endereço, bairro, condomínio, idade, sexo.

Figura 1- Banco de dados

	A	B	C	D	E	F	G	H	I	J
	Data da emissão	Data do vencimento	Data do pagamento	DIAS DE ATRASO	Pago em dia	Valor recebido	Bairro	Condomínio	idade	Sexo
339	28/12/2021	01/10/2022	02/01/2022	272	VERDADEIRO	R\$ 81,90	Jabutiana	Bossa Nova Condominio Clube	26	Masculino
340	07/12/2021	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 64,90	Jabutiana	Condominio Exclusive Jabutiana	70	Feminino
341	25/01/2022	02/05/2022	02/01/2022	120	VERDADEIRO	R\$ 62,91	Jabutiana	Condominio Encantos do Bosque	70	Feminino
342	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 79,90	Luzia	Repletto Condominio Clube	70	Feminino
343	14/01/2022	25/01/2022	02/01/2022	23	VERDADEIRO	R\$ 81,60	Jabutiana	Viamonte Condominio Clube	68	Masculino
344	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 79,90	Jabutiana	Condominio Privilege Residence	66	Masculino
345	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 64,90	Jabutiana	Condominio Privilege Residence	66	Masculino
346	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 99,90	Jabutiana	Life Jabotiana	65	Masculino
347	14/01/2022	25/01/2022	02/01/2022	23	VERDADEIRO	R\$ 102,06	Zona de Expansão	Condominio Melicio Machado	65	Masculino
348	30/11/2021	15/12/2021	02/01/2022	-18	FALSO	R\$ 82,42	Jabutiana	Life Jabotiana	62	Masculino
349	05/01/2022	15/01/2022	02/01/2022	13	VERDADEIRO	R\$ 81,80	Jabutiana	Life Jabotiana	62	Masculino
350	06/01/2022	20/01/2022	02/01/2022	18	VERDADEIRO	R\$ 66,40	Inacio Barbosa	Residencial Alameda dos Pássaros	62	Feminino
351	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 79,90	Inacio Barbosa	Residencial Alameda dos Pássaros	61	Masculino
352	06/01/2022	20/01/2022	02/01/2022	18	VERDADEIRO	R\$ 81,70	Inacio Barbosa	Condominio Alpha Park	60	Masculino
353	25/01/2022	02/05/2022	02/01/2022	120	VERDADEIRO	R\$ 71,91	Inacio Barbosa	Condominio Alameda dos Ipes	60	Masculino
354	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 132,90	Jabutiana	Condominio Natura Ville	59	Masculino
355	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 129,90	Siqueira Campos	Vivendas De Aracaju	59	Feminino
356	06/01/2022	20/01/2022	02/01/2022	18	VERDADEIRO	R\$ 81,70	Inacio Barbosa	Condominio Alpha Park	59	Masculino
357	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 99,90	Jabutiana	Bossa Nova Condominio Clube	57	Feminino
358	19/01/2022	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 79,90	Jabutiana	Condominio Natura Ville	56	Masculino
359	17/12/2021	30/12/2021	02/01/2022	-3	FALSO	R\$ 66,82	Jabutiana	Condominio Laguna Mar	56	Feminino
360	28/12/2021	01/10/2022	02/01/2022	272	VERDADEIRO	R\$ 66,60	Jabutiana	Condominio Specialle Residenza	55	Feminino
361	20/12/2021	30/01/2022	02/01/2022	28	VERDADEIRO	R\$ 64,90	Jabutiana	Life Jabotiana	54	Masculino
362	25/01/2022	02/05/2022	02/01/2022	120	VERDADEIRO	R\$ 89,91	Farolândia	Condominio Le Soleil	54	Masculino
363	10/12/2021	14/12/2021	02/01/2022	-19	FALSO	R\$ 34,14	Farolândia	Vitta Condominio Club	53	Masculino
364	28/12/2021	01/10/2022	02/01/2022	272	VERDADEIRO	R\$ 143,19	Farolândia	Vitta Condominio Club	53	Masculino
365	25/01/2022	02/05/2022	02/01/2022	120	VERDADEIRO	R\$ 85,41	Inacio Barbosa	Condominio Alameda das Águas	51	feminino
366	25/01/2022	02/05/2022	02/01/2022	120	VERDADEIRO	R\$ 81,70	Jabutiana	Condominio Natura Ville	49	Feminino

Fonte: Elaborado pelo autor

Por questões de proteção de dados e pela não necessidade de utilização no estudo, as colunas nome do cliente, CPF, e endereço foram excluídas.

3.3 VARIÁVEIS UTILIZADAS

Esta subseção tem como objetivo explicar cada variável utilizada no banco de dados. As variáveis utilizadas no estudo foram selecionadas com base na literatura sobre inadimplência e na disponibilidade de dados no sistema IXC Soft, o quadro 4 apresenta quais foram as variáveis e sua descrição.

Quadro 4 – variáveis do banco de dados

VARIÁVEL	DESCRIÇÃO
DATA DE EMISSÃO	Data em que a fatura do cliente foi gerada.
DATA DE VENCIMENTO	Data de vencimento do boleto do cliente, se pago até esta data o cliente está adimplente, passando disso se torna inadimplente.
DATA DE PAGAMENTO	Data em que o cliente efetuou o pagamento do boleto.
DIAS DE ATRASO	Coluna criada pelo autor, representa a diferença entre a data de pagamento e a data de vencimento: se o valor for menor que 0, indica quantos dias o cliente pagou o boleto antecipadamente; se for igual a 0, significa que o cliente efetuou o pagamento na data de vencimento; e, se for maior que 0, indica quantos dias o cliente atrasou o pagamento, caracterizando a inadimplência.
PAGO EM DIA	coluna criada pelo autor indica a situação em que o cliente pagou sua fatura, com base nos dias de atraso. Se o resultado for 'Verdadeiro', significa que o cliente pagou com adimplência, o que indica que o valor dos dias de atraso é maior ou igual a zero. Por outro lado, se o resultado for 'Falso', isso significa que o cliente pagou com inadimplência, o que implica que o valor dos dias de atraso é menor que zero.

VALOR RECEBIDO	Valor recebido do boleto pago pelo cliente.
BAIRRO	Bairro onde se localiza a residência do cliente.
CONDOMÍNIO	Condomínio onde se localiza a residência do cliente.
IDADE	Idade do cliente (em anos).
SEXO	Gênero do cliente da conta (masculino ou feminino).

Fonte: Elaborado pelo autor

3.4 ANÁLISE DOS DADOS E HIPÓTESES

Para Gil (2010), Hipóteses são componentes fundamentais na pesquisa científica, desempenhando um papel crucial no processo de investigação e no desenvolvimento de estudos rigorosos. Elas são afirmações cuidadosamente elaboradas que têm o propósito de descrever e testar relações entre variáveis, auxiliando os pesquisadores a avançarem em direção a uma compreensão mais profunda de fenômenos específicos.

Com base na literatura sobre inadimplência e na análise exploratória dos dados disponíveis no sistema IXC Soft, foram formuladas as seguintes hipóteses:

1. O valor pago mensalmente tem relação com a inadimplência, clientes que pagam valores mais altos podem ter um risco maior de atraso.
2. O gênero influencia na inadimplência, diferenças entre clientes do sexo masculino e feminino na taxa de adimplência e inadimplência.
3. A idade influencia na inadimplência, clientes mais jovens tendem a apresentar maior taxa de atraso nos pagamentos.
4. A taxa de condomínio impacta a inadimplência, valores mais elevados podem dificultar o pagamento.

3.4.1 Metodologia para Teste de Hipóteses

Para testar essas hipóteses, foi empregado o método de correlação de Pearson e Modelo T Student, por meio do software R studio, estes métodos oferecem *insights* valiosos sobre as relações lineares entre variáveis.

A análise foi conduzida considerando um nível de significância de 0,05, garantindo a confiabilidade dos resultados. A estatística descritiva foi utilizada para entender a distribuição dos dados antes de aplicar os testes.

4. ANÁLISE E DISCUSSÃO DOS RESULTADOS

Este capítulo tem como propósito primordial apresentar e discutir os resultados obtidos a partir da aplicação dos métodos e procedimentos descritos na seção de metodologia, nesse momento as informações coletadas são interpretadas à luz dos objetivos propostos e da fundamentação teórica abordada.

4.1 ESTATÍSTICA DESCRITIVA

O quadro a seguir fornece uma análise detalhada da inadimplência e inadimplência em relação a diferentes variáveis. Ela mostra a distribuição de adimplentes e inadimplentes dos pagamentos realizados pelos clientes, bem como a taxa de inadimplência e inadimplência e os totais das frequências relativas em várias categorias.

Para o cálculo das taxas será utilizado:

TAXA DE ADIMPLÊNCIA

$$= ((total\ de\ pagos\ em\ dia) / (total\ de\ pagamentos\ recebidos)) . 100$$

TAXA DE INADIMPLÊNCIA

$$= ((total\ de\ pagos\ em\ atraso) / (total\ de\ pagamentos\ recebidos)) . 100$$

Quadro 5 - Distribuição das frequências absolutas e relativas das variáveis da base de dados

VARIÁVEIS		ADIMPLÊNCIA		INADIMPLÊNCIA		TOTAL
		Freq. Absoluta	Taxa de Adimp. %	Freq. Absoluta	Taxa de Inadimp. %	Freq. Absoluta
Total		1.682	45,00	2.056	55,00	3.738
Gênero	Masculino	944	45,65	1.124	54,35	2.068
	Feminino	738	44,19	932	55,81	1.670
Idade	18 a 25	105	39,47	161	60,53	266
	26 a 35	510	44,46	637	55,54	1.147
	36 a 45	647	46,38	748	53,62	1.395
	46 a 55	228	41,83	317	58,17	545
	56 a 65	126	51,43	119	48,57	245
	66 a 75	58	48,33	62	51,67	120
	76+	8	40,00	12	60,00	20
Data de Vencimento	5	341	52,95	303	47,05	644
	10	544	50,89	525	49,11	1.069

	15	401	45,72	476	54,28	877
	20	118	31,38	258	68,62	376
	25	193	39,96	290	60,04	483
	30	85	29,41	204	70,59	289
Valor da Mensalidade	50 a 70,99	403	44,00	513	56,00	916
	71 a 84,99	461	46,10	539	53,90	1.000
	85 a 115,99	712	45,91	839	54,09	1.551
	116+	106	39,11	165	60,89	271
Localidade	zona oeste	1.044	44,71	1.291	55,29	2.335
	zona sul	601	45,74	713	54,26	1.314
	centro	31	46,27	36	53,73	67
	expansão	6	27,27	16	72,73	22

Fonte: Base de dados da pesquisa

O quadro 6 é uma tabela que apresenta algumas estatísticas descritivas para as variáveis quantitativas da base de dados

Quadro 6 – Estatística descritiva

SITUAÇÃO	INDICADOR	VALOR mens	DIAS	IDADE
ADIMPLÊNCIA	MÉDIA	82,57218904	-3,42698	39,3478
	MEDIANA	83,393	-3	38
	MODA	90,7425	-3	36
	DESVIO PADRÃO	19,86313146	2,526217	11,43035
	MÁXIMO	254,1516667	0	86
	MINIMO	60,65	-22	18
INADIMPLÊNCIA	MÉDIA	85,94040239	8,619347	38,6644
	MEDIANA	83,9	5,2	37
	MODA	90,7425	1	37
	DESVIO PADRÃO	21,33555708	9,011845	11,23001
	MÁXIMO	276,0081818	68,66667	85
	MINIMO	50,99	0,083333	18
TOTAL	MÉDIA	84,42479649	3,198823	38,97191
	MEDIANA	83,77025641	0,75	37
	MODA	90,7425	-3	36
	DESVIO PADRÃO	20,75103392	9,135389	11,32418
	MÁXIMO	276,0081818	68,66667	86
	MINIMO	50,99	-22	18

Fonte: Elaborado Pelo autor

4.1.1 Análise da estatística descritiva

Analisando a estatística descritiva do quadro 6, pode-se notar algumas informações consideráveis:

Principais Tópicos Relevantes da estatística descritiva:

1. Valor Mensal:
 - A média de valor recebido é ligeiramente maior para inadimplentes (R\$85,94) do que para adimplentes (R\$82,57).
2. Dias de Pagamento:
 - Adimplentes pagam, em média, 3,43 dias antes do vencimento, enquanto inadimplentes pagam 8,62 dias após o vencimento.
 - A variação nos dias de pagamento é muito maior para inadimplentes (desvio padrão de 9,01) em comparação com adimplentes (desvio padrão de 2,53).
3. Idade:
 - A idade média dos adimplentes (39,35 anos) é ligeiramente superior à dos inadimplentes (38,66 anos), embora a diferença seja pequena.
4. Desvios Padrão e Extremos:
 - A variação no valor recebido é maior entre inadimplentes (desvio padrão de R\$21,34) do que entre adimplentes (desvio padrão de R\$19,86).

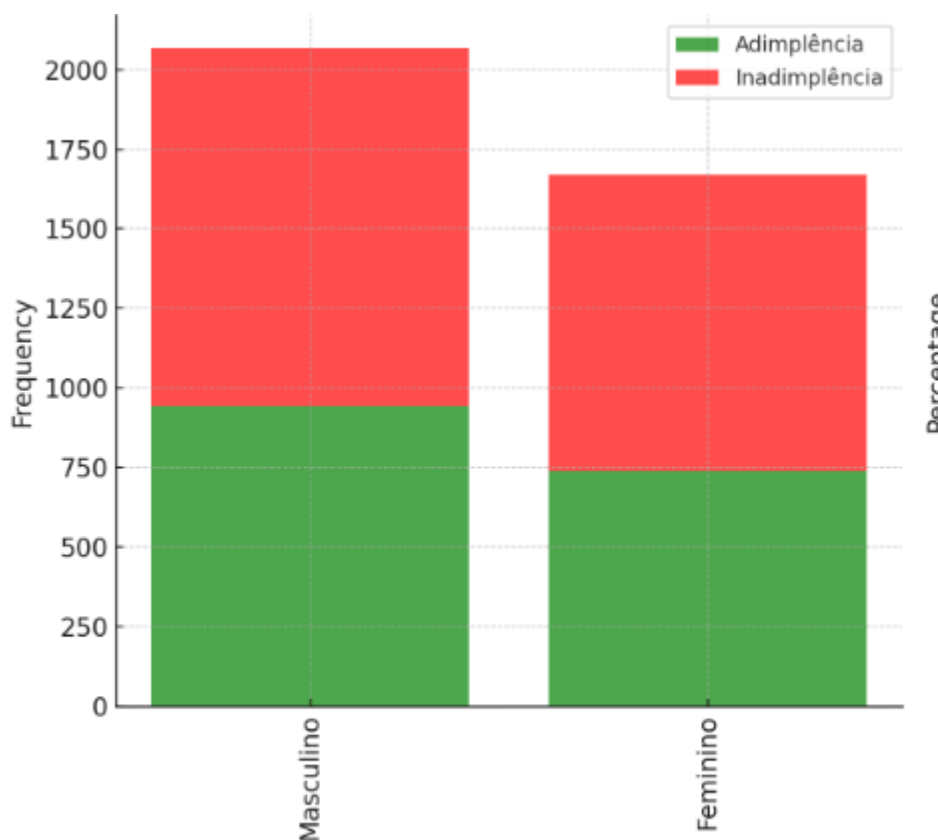
4.1.2 Tópicos Relevantes da estatística descritiva e da distribuição de frequência

Pode-se notar alguns fatores interessantes no quadro 5 nas variáveis que serão analisadas individualmente

a) Gênero :

As taxas de adimplência e inadimplência são semelhantes entre homens e mulheres, com uma ligeira predominância de inadimplência entre ambos os gêneros.

Figura 2 – Adimplência e inadimplência x gênero



Fonte: Elaborado pelo autor

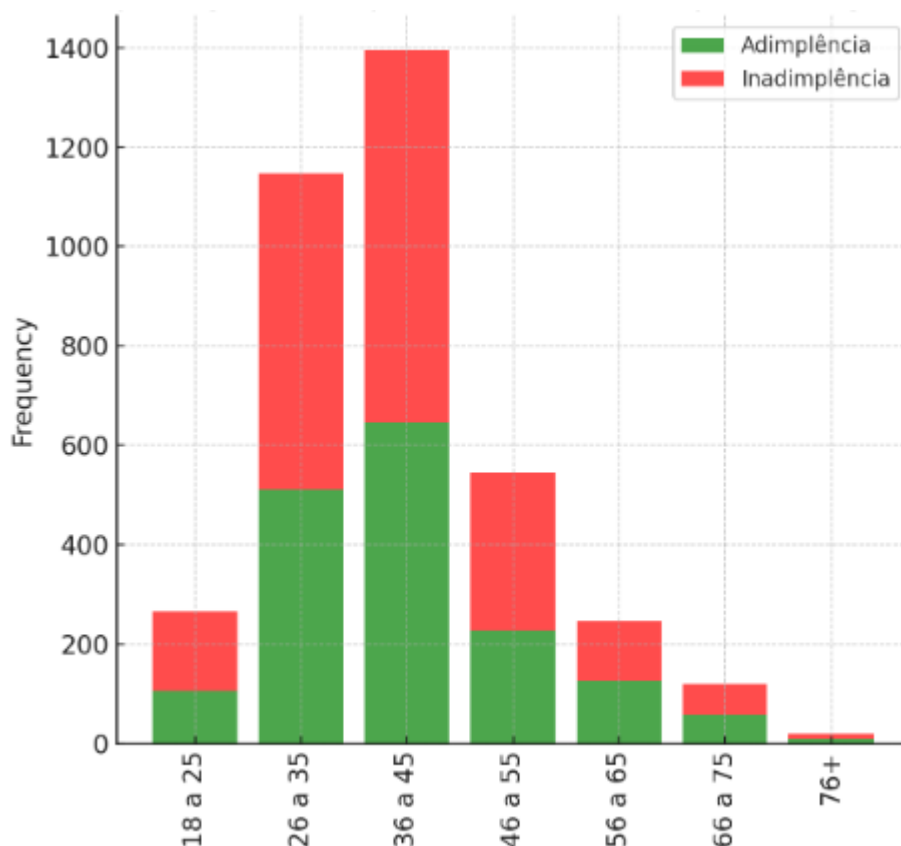
b) Idade:

A faixa etária de 18 a 25 anos apresenta a maior taxa de inadimplência (60,53%), enquanto a faixa de 56 a 65 anos tem a maior taxa de adimplência (51,43%).

A alta inadimplência entre os jovens pode indicar a necessidade de educação financeira específica para essa faixa etária.

A faixa etária de 76 anos ou mais também apresenta uma alta taxa de inadimplência (60%), sugerindo possíveis dificuldades financeiras ou de gestão de pagamentos entre os mais velhos.

Figura 3 - Adimplência e inadimplência por faixa etária



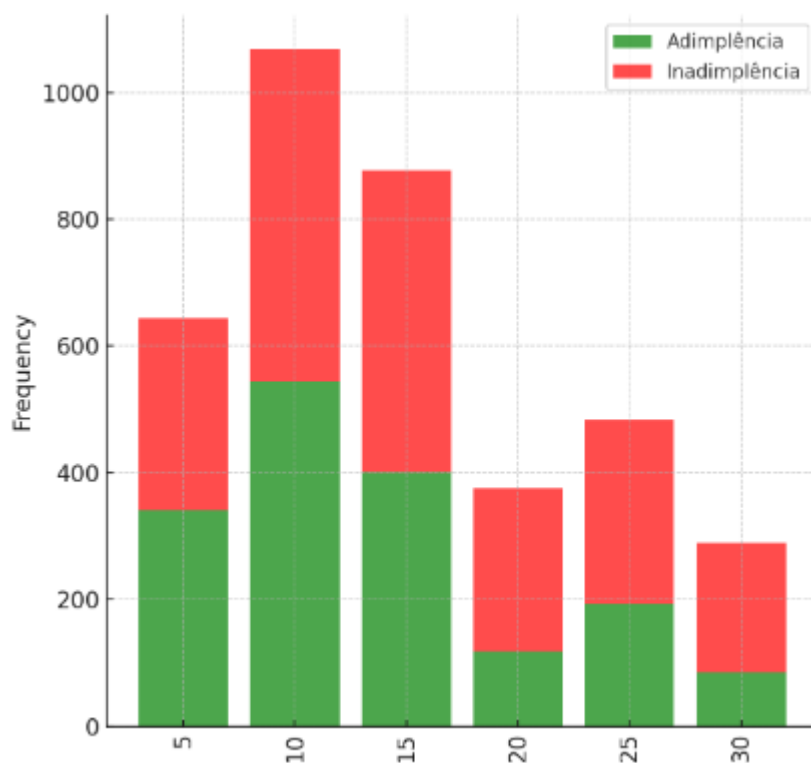
Fonte: elaborado pelo autor

c) Data de vencimento :

Dias de vencimento no início do mês (dia 5 e dia 10) têm taxas de adimplência mais altas (52,95% e 50,89%).

Dias de vencimento mais tardios (dias 20, 25, e 30) apresentam maiores taxas de inadimplência (68,62%, 60,04%, e 70,59%, respectivamente) ou seja conforme o dia do vencimento se aproxima do final do mês a proporção de inadimplentes aumenta, Isso pode indicar que os clientes têm mais dificuldades em gerenciar pagamentos no final do mês

Figura 4 – Distribuição por data de vencimento

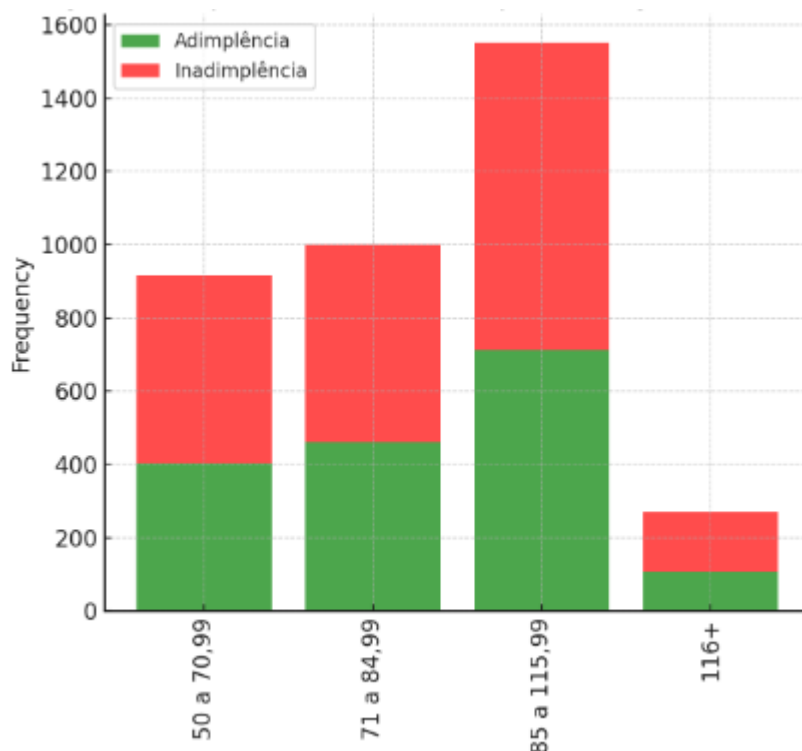


Fonte: elaborado pelo autor

d) Valor da Mensalidade:

Mensalidades na faixa de R\$50 a R\$70,99 e R\$116+ têm taxas de inadimplência mais altas (56,00% e 60,89%, respectivamente). Valores mais altos podem ser mais difíceis de gerenciar, contribuindo para a inadimplência.

As faixas de mensalidade entre R\$71 a R\$84,99 e R\$85 a R\$115,99 apresentam taxas de adimplência ligeiramente mais equilibradas.

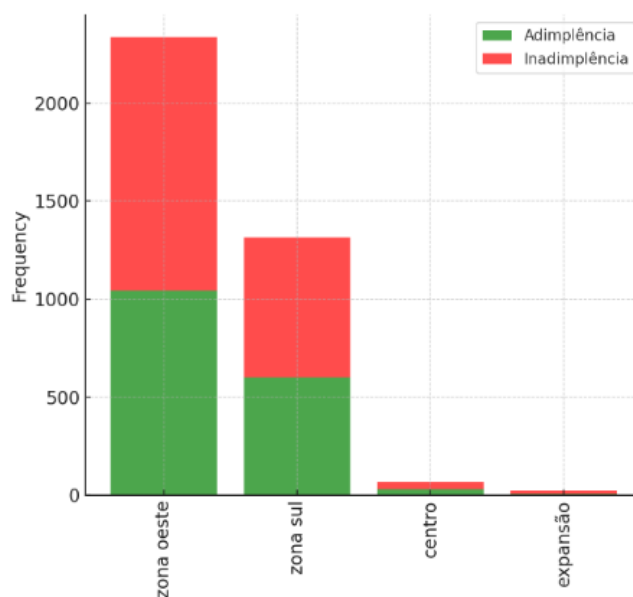
Figura 5 – Distribuição por mensalidade

Fonte: elaborado pelo autor

e) Localidade:

A zona oeste tem a maior quantidade absoluta de inadimplentes (1.291), seguida pela zona sul (713).

A taxa de inadimplência é mais alta na expansão (72,73%), possivelmente devido a fatores socioeconômicos ou acesso a serviços financeiros.

Figura 6 - Distribuição por localidade

Fonte: elaborado pelo autor

Após a distribuição de frequência das variáveis de adimplência e inadimplência, podemos perceber padrões significativos em cada uma das categorias analisadas. Com base nos gráficos apresentados e nos tópicos destacados, é possível identificar fatores que podem influenciar os comportamentos financeiros dos clientes em diferentes contextos.

4.2 HIPOTEESES

Este capítulo tem como objetivo realizar testes de hipóteses para investigar diferentes relações e padrões nos seus dados.

4.2.1 Situação do cliente por valor da mensalidade.

Nesta etapa do trabalho, foi realizada uma análise estatística inferencial para testar a hipótese de que o valor recebido pelos clientes afeta a sua adimplência ou inadimplência

Foi executado o teste t de Student diretamente no software Rstudio comparando as médias dos valores recebidos entre os clientes adimplentes e inadimplentes.

O teste t de Student foi escolhido porque é uma técnica estatística apropriada para comparar as médias de duas populações independentes, determinando se existe uma diferença estatisticamente significativa entre elas. No presente estudo, a variável de interesse é o valor recebido pelos clientes adimplentes e inadimplentes. Como se busca verificar se há uma diferença média significativa entre os dois grupos, o teste t é adequado, pois permite avaliar se a diferença observada nas médias pode ser atribuída ao acaso ou se representa uma tendência real nos dados analisados (WELCH, 1947; DEVORE, 2011).

Figura 7 – Fórmula do teste t de student

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Fonte: (Welch, 1947; Devore, 2011).

Onde:

$\bar{X}_1$: Média dos valores recebidos pelos clientes adimplentes

$\bar{X}_2$: Média dos valores recebidos pelos clientes inadimplentes

S_1^2 , S_2^2 : são as variâncias das amostras 1 e 2, respectivamente.

n_1 , n_2 : são os tamanhos das amostras 1 e 2, respectivamente.

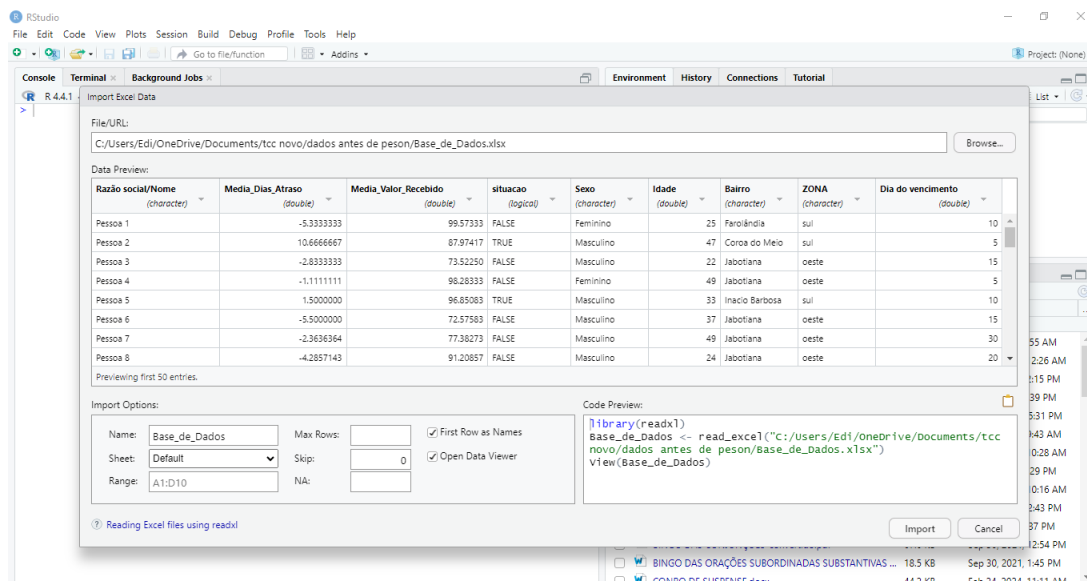
Hipótese nula (H_0): As médias dos dois grupos são iguais. Ou seja, não há diferença entre as médias.

Hipótese alternativa (H_1): As médias dos dois grupos são diferentes. Ou seja, há uma diferença entre as médias dos grupos adimplente e inadimplente.

O teste segue os seguintes passos:

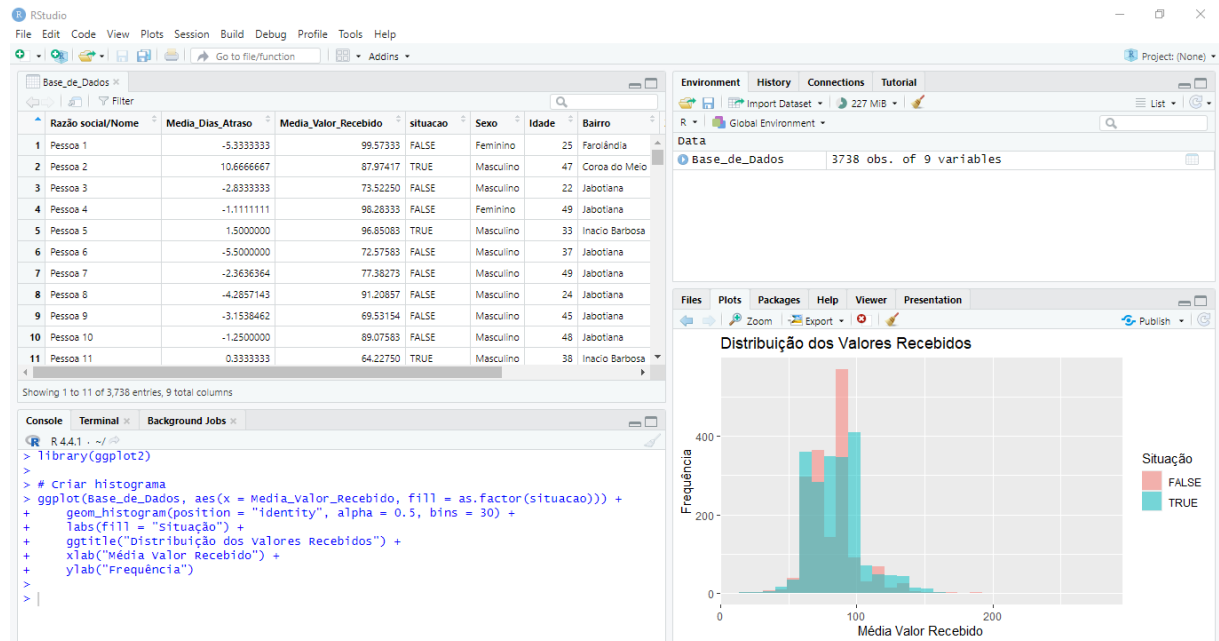
Figura 8 - Passo a passo para teste t student para valor da mensalidade no R

A) Importação da Base de dados no software



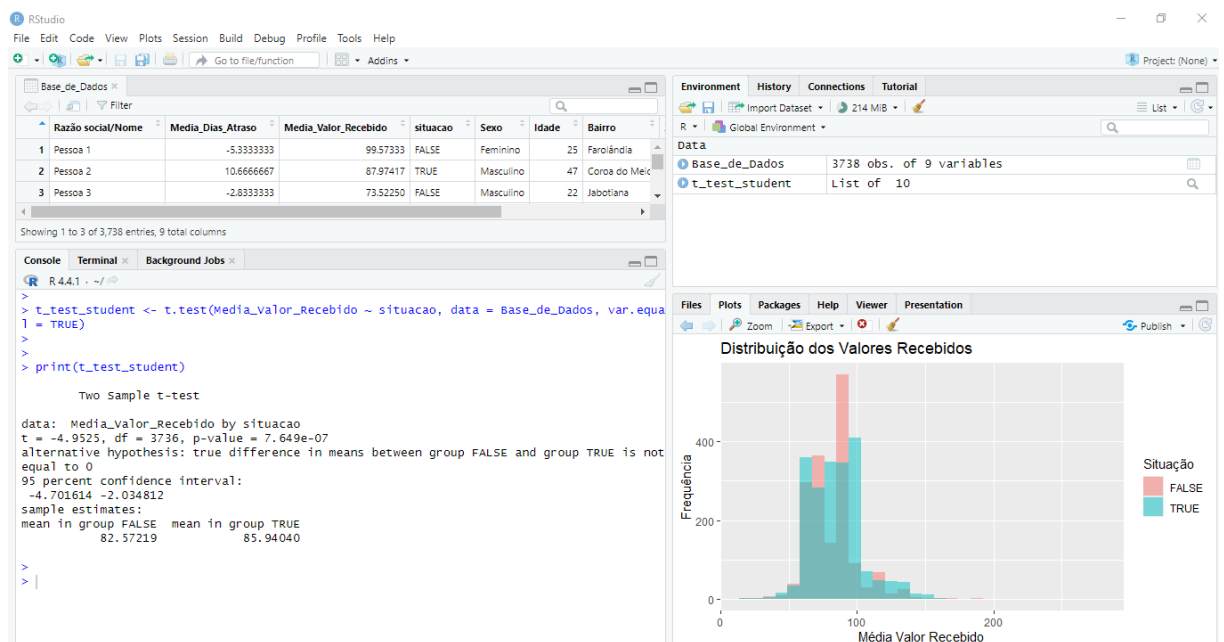
Fonte: elaborado pelo autor

B) Criar um histograma para visualizar a distribuição dos valores recebidos



Fonte: elaborado pelo autor

C) Execução do teste no software



Fonte: elaborado pelo autor

4.2.1.1 interpretação dos resultados fornecidos

Resultados do Teste t

1. **Estatística t (t-value):** -4.9525
2. **Graus de liberdade (df):** 3736
3. **Valor-p (p-value):** 0.0000007649
4. **Intervalo de confiança de 95% para a diferença de médias:** -4.701614 a -2.034812
5. **Estimativas das médias dos grupos:**
 - Média do grupo adimplente (FALSE): 82.57219
 - Média do grupo inadimplente (TRUE): 85.94040

Interpretação

1. Valor-p (p-value)

O valor-p é 0.0000007649, que é muito menor do que o nível de significância comum (0.05). Isso indica que há uma diferença estatisticamente significativa entre os valores médios recebidos dos clientes adimplentes e inadimplentes. De acordo com Devore (2011), o valor-p é a probabilidade de se obter um resultado tão extremo quanto o observado, assumindo que a hipótese nula seja verdadeira. Valores-p baixos, como neste caso, sugerem a rejeição da hipótese nula em favor da hipótese alternativa.

2. Estatística t (t-value)

O valor t de -4.9525 indica que a média do grupo adimplente é significativamente menor que a média do grupo inadimplente. A estatística t é utilizada para avaliar a diferença entre as médias de dois grupos, considerando a variação dos dados (Montgomery; Runger, 2014; Devore, 2011). Um valor t negativo, como neste caso, indica que a média do grupo adimplente é inferior à média do grupo inadimplente.

3. Intervalo de confiança

O intervalo de confiança de 95% para a diferença de médias varia de -4.692129 a -2.044298. Como o intervalo não inclui 0, isso reforça a conclusão de que há uma diferença significativa entre os dois grupos. Segundo Lehmann e Romano (2005), o intervalo de confiança oferece uma faixa dentro da qual a verdadeira diferença de

médias provavelmente se encontra, com um determinado nível de confiança. Um intervalo que não inclui 0 sugere que a diferença observada é estatisticamente significativa.

4. Médias dos grupos

A média do valor recebido para os clientes adimplentes é 82.57219, enquanto a média do valor recebido para os clientes inadimplentes é 85.94040. A diferença entre as médias é negativa, indicando que, em média, os clientes adimplentes pagam menos que os inadimplentes. Como destacado por Moore, McCabe e Craig (2012), comparar as médias de dois grupos é uma maneira comum de avaliar diferenças entre eles. A diferença negativa observada sugere uma tendência nos dados, onde os inadimplentes tendem a receber valores maiores.

4.2.1.2 Decisão

Rejeita-se a hipótese nula (H_0).

Justificativa

Valor-p:

O valor-p sendo extremamente menor do que o nível de significância (0.05). Um valor-p tão pequeno indica que a probabilidade de uma diferença tão extrema entre as médias dos dois grupos apenas por acaso é muito baixa.

Intervalo de Confiança:

O intervalo de confiança de 95% para a diferença de médias varia de -4.701614 a -2.034812. Este intervalo não inclui 0, o que reforça a conclusão de que há uma diferença significativa entre as médias dos dois grupos.

Médias dos Grupos:

A média dos valores recebidos pelos clientes adimplentes é 82.57219, enquanto a média dos valores recebidos pelos clientes inadimplentes é 85.94040. A diferença entre essas médias é significativa e indica que, em média, os clientes inadimplentes recebem valores maiores que os adimplentes o que é suficiente para rejeitar a hipótese nula.

4.2.1.3 Conclusão

Os resultados da análise indicam que há uma relação significativa entre o valor recebido e a situação de adimplência dos clientes. Especificamente, clientes que pagam valores maiores tendem a ser mais inadimplentes. Isso sugere que o valor pode ser um fator influente na probabilidade de inadimplência dos clientes.

4.2.2 Situação do cliente por gênero

Nesta etapa do trabalho, realizou-se uma análise estatística inferencial para testar a hipótese de que o gênero dos clientes afeta a sua adimplência ou inadimplência.

Hipóteses

Hipótese Nula (H_0): Não há associação significativa entre o gênero dos clientes e a situação de pagamento. Em outras palavras, a distribuição dos clientes adimplentes e inadimplentes é independente do gênero.

Hipótese Alternativa (H_1): Existe uma associação significativa entre o gênero dos clientes e a situação de pagamento. Ou seja, a distribuição dos clientes adimplentes e inadimplentes está relacionada ao gênero.

Para testar essa relação, foi utilizado o teste Qui-quadrado de Independência, que é uma técnica estatística amplamente empregada para avaliar a associação entre variáveis categóricas. Esse teste verifica se a distribuição observada dos dados difere da esperada sob a hipótese de independência, permitindo identificar possíveis relações entre as variáveis analisadas (Tavares, 1996). A análise foi conduzida diretamente no software RStudio, utilizando as frequências observadas de clientes adimplentes e inadimplentes por gênero, possibilitando verificar se há uma relação estatisticamente significativa entre o gênero dos clientes e sua situação de pagamento.

Fórmula do teste Qui-quadrado

Figura 9 – Fórmula do teste Qui-quadrado

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Fonte: (Agresti, 2007)

onde:

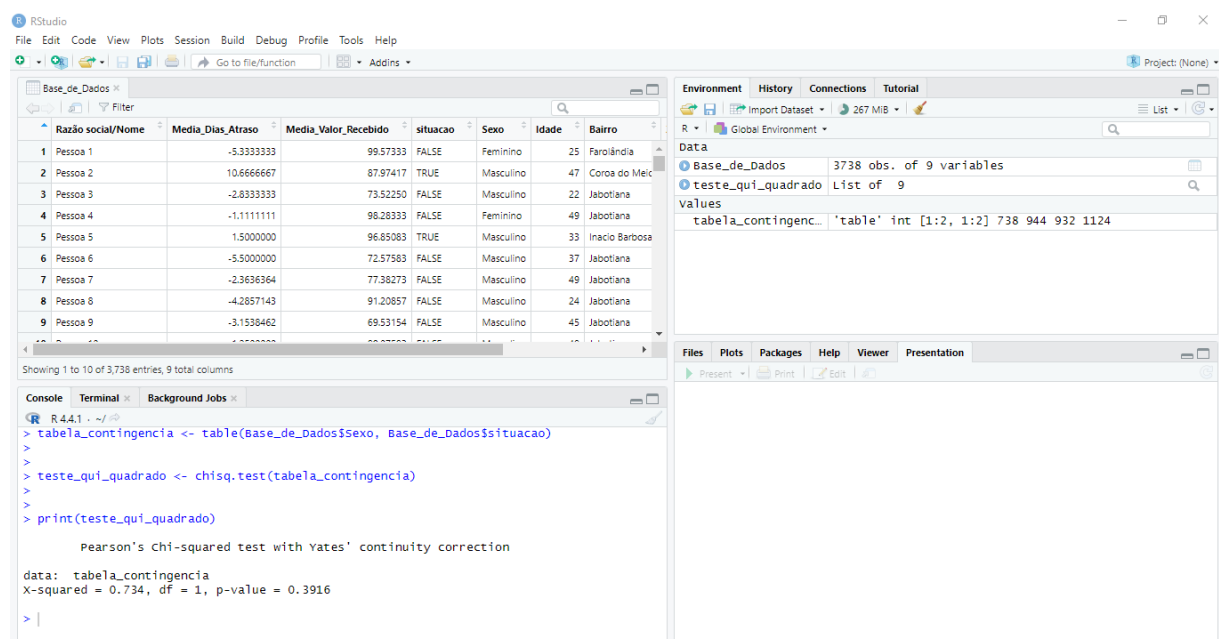
χ^2 : Estatística do Qui-quadrado

O_i : Frequência observada em cada categoria

E_i : Frequência esperada em cada categoria

Com a base de dados já importada no software devido ao teste anterior, será realizada diretamente à execução

Figura 10 – Teste qui-quadrado para gênero no R



Fonte: elaborado pelo autor

4.2.2.1 Interpretação dos resultados fornecidos

Resultados do Teste

- Estatística do Qui-quadrado (X^2): 0.734
- Graus de liberdade (df): 1
- Valor-p (p-value): 0.3916

Interpretação

Valor-p (p-value):

O valor-p de 0.3916 é maior que o nível de significância padrão de 0,05. Isso significa que a probabilidade de observar a diferença entre os grupos de gênero em termos de adimplência, apenas por acaso, é alta. Portanto, não há evidências estatisticamente significativas para rejeitar a hipótese nula (H_0). Segundo Montgomery e Runger (2014), o valor-p é uma medida da evidência contra a hipótese nula; valores altos sugerem que as diferenças observadas provavelmente ocorreram devido ao acaso.

Estatística do Qui-quadrado (X^2):

O valor de 0.734 sugere que, mesmo se houver alguma associação entre gênero e adimplência, ela é muito fraca. Como observado por Agresti (2007), a estatística do Qui-quadrado é usada para avaliar a associação entre variáveis categóricas; valores baixos indicam uma associação fraca ou inexistente entre as variáveis testadas.

4.2.2.2 Decisão

A hipótese nula (H_0), não foi rejeitada

Justificativa:

Dado que o valor-p de 0.3916 é significativamente maior que o nível de significância de 0,05, e considerando que a estatística do Qui-quadrado foi de 0.734, não há evidências estatisticamente significativas para rejeitar a hipótese nula. Isso sugere que, o gênero dos clientes não afeta de forma significativa a probabilidade de eles serem adimplentes ou inadimplentes

4.2.3 Situação do cliente por idade

Nesta etapa foi testada a hipótese de que a idade dos clientes tem impacto adimplência ou inadimplência

Hipóteses

Hipótese Nula (H0): Não há correlação entre as duas variáveis (idade e a situação de pagamentos dos clientes) $r=0$

Hipótese Alternativa (H1): Há correlação entre as duas variáveis (idade e a situação de pagamentos dos clientes) $r \neq 0$

Foi executado o teste da correlação de Pearson diretamente no software Rstudio. A correlação de Pearson é uma medida estatística que avalia a força e a direção da relação linear entre duas variáveis contínuas, um valor de “r” próximo de 0 indica uma correlação fraca ou inexistente, enquanto valores próximos de +1 ou -1 indicam uma correlação forte (DEVORE, 2011).

Figura 11 – Fórmula da correlação de Pearson

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Fonte: (Pearson, 1896; Devore, 2011)

Onde:

r: é o coeficiente de correlação de Pearson

X: é a variável idade

Y: é a variável média de dias

X_i : são os valores individuais da variável idade

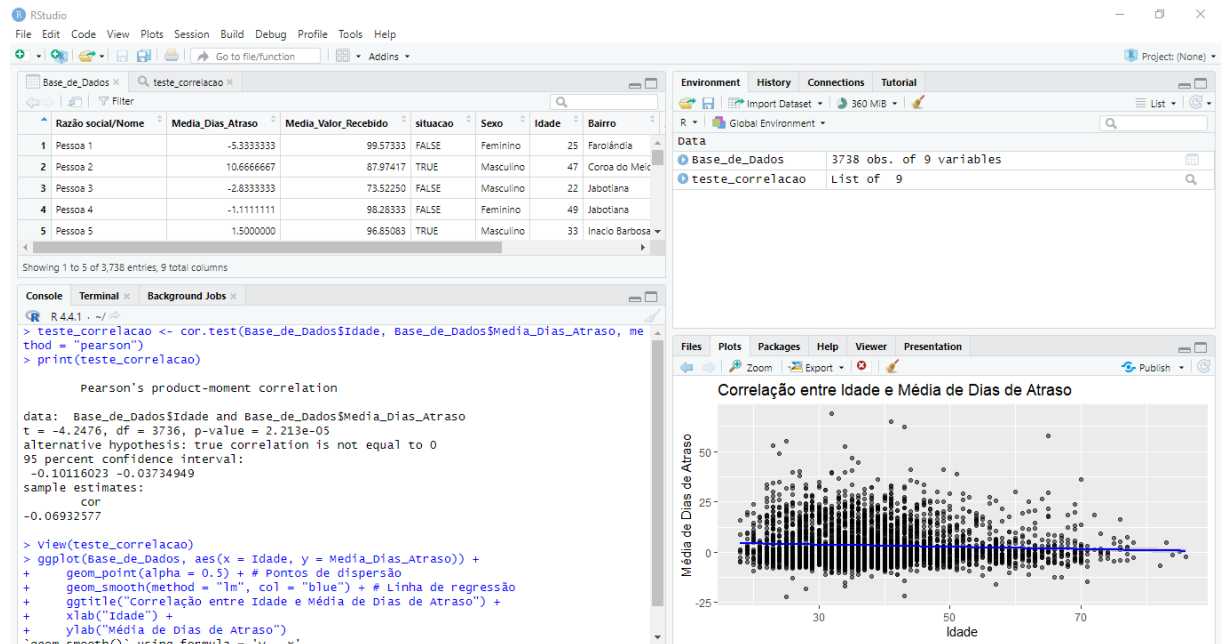
Y_i : são os valores individuais da variável média de dias

$\bar{x}$: é a média da variável idade

$\bar{y}$: é a média da variável média de dias

Com a base de dados já importada no software devido aos testes anteriores, foi realizada diretamente à execução.

Figura 12 – teste de correlação de Pearson para idade no R



Fonte: elaborado pelo autor

4.2.3.1 Interpretação dos resultados fornecidos

Resultados do Teste:

- Estatística t: -4.2476
- Graus de liberdade (df): 3736
- Valor de r (correlação): -0.0693
- Valor-p: 0.00002213

Interpretação

O valor de $r = -0.0693$ indica que há uma correlação negativa fraca entre a idade dos clientes e a média de dias de atraso ou adiantamento. Isso sugere que, à medida que a idade aumenta, há uma leve tendência dos clientes a atrasarem menos os pagamentos ou a pagarem com um pouco mais de antecedência, mas essa relação é fraca (Devore, 2011). Porém valor-p de 0.00002213, que é extremamente pequeno, indica que a correlação observada é estatisticamente significativa. Segundo Montgomery e Runger (2014), um valor-p tão pequeno sugere que a probabilidade de a correlação observada ter ocorrido por acaso é muito baixa, o que leva a rejeitar a hipótese nula.

4.2.3.2 Decisão

Dado o $r \neq 0$ e o valor-p de 0.00002213, que é muito menor do que o nível de significância comum de 0,05, rejeita-se a hipótese nula (H_0), concluindo que há uma correlação estatisticamente significativa entre a idade dos clientes e a situação de pagamentos.

4.3 Situação do cliente por taxa de condomínio

Nesta etapa do estudo, tem-se como objetivo principal explorar e compreender a relação entre a taxa de condomínio e como essa interação influencia os comportamentos de adimplência ou inadimplência.

A interação entre a renda dos clientes e suas obrigações financeiras constitui um tema central no estudo das dinâmicas de adimplência e inadimplência, particularmente no que concerne às despesas fixas mensais, como as taxas de condomínio. Embora representem uma parcela constante do orçamento doméstico, podem ter um impacto significativo sobre a capacidade financeira dos consumidores, influenciando, assim, suas decisões de pagamento.

Com base na literatura existente, que frequentemente associa a inadimplência a fatores como variações na renda familiar (Dyner; Skinner; Zeldes, 2000) e alterações nas condições econômicas (Mian, Rao, & Sufi, 2013), esta pesquisa analisou o impacto de variáveis financeiras sobre a adimplência, incluindo a taxa de condomínio. No entanto, não foram coletados dados diretos sobre a renda dos clientes, impossibilitando a verificação de modificações nesse aspecto ao longo do período estudado. Dessa forma, a relação entre inadimplência e renda foi considerada apenas com base nos referenciais teóricos disponíveis, enquanto a taxa de condomínio foi incorporada à análise como uma variável potencialmente crítica no perfil dos clientes.

4.3.1 Coleta dos dados

Para coletar os dados com as taxas de condomínio, optou-se por uma abordagem mista, que incluiu a colaboração com imobiliárias, a análise de anúncios imobiliários online e o contato direto com as portarias dos condomínios.

Sendo assim foi adicionada mais uma coluna chamada “taxa de condomínio”, na Data Lake original, a técnica estatística utilizada foi a correlação de Pearson calculado no software R Studio, para amostras independentes. Este teste visa

comparar as taxas de condomínio e média de dias em que os clientes costumam adiantar ou atrasar seus pagamentos, permitindo avaliar a existência de uma diferença estatística significativa entre eles.

4.3.2 Formulação do teste

Hipótese:

Hipótese Nula (H_0): Não há correlação entre as duas variáveis (taxa de condomínio e a situação de pagamentos dos clientes) $r=0$

Hipótese Alternativa (H_1): Não há correlação entre as duas variáveis (taxa de condomínio e a situação de pagamentos dos clientes) $r \neq 0$

Utilizando a mesma fórmula apresentada na seção 4.2.3 (Fórmula 1), obtêm-se as seguintes variáveis:

Onde:

r : é o coeficiente de correlação de Pearson

X : é a variável taxa de condomínio

Y : é a variável média de dias

X_i : são os valores individuais da variável taxa de condomínio

Y_i : são os valores individuais da variável média de dias

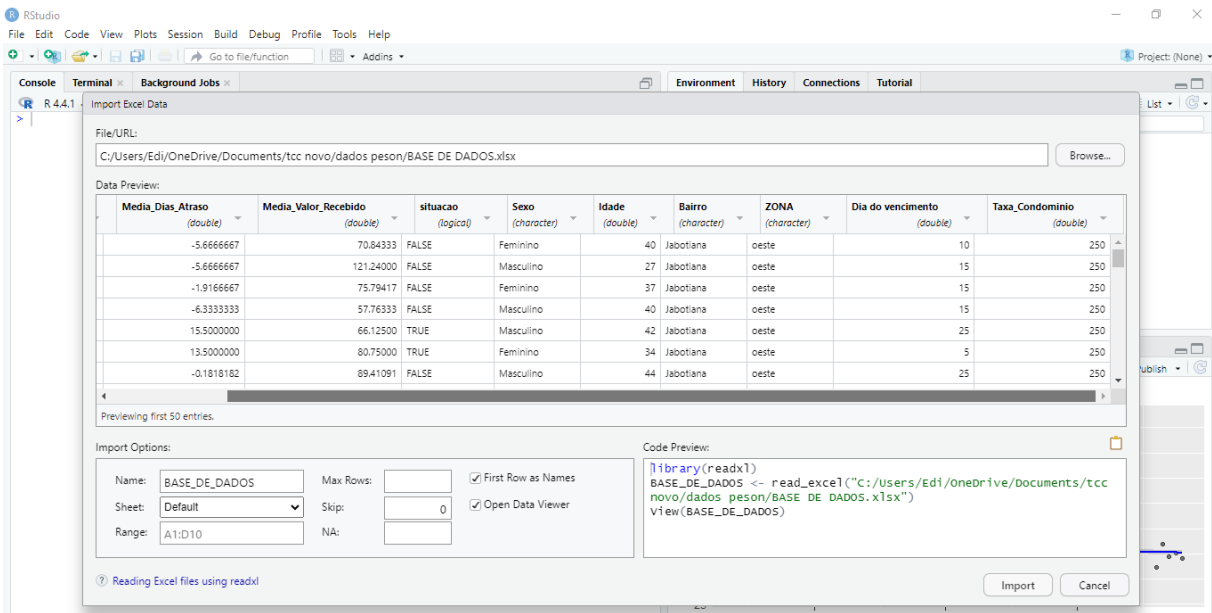
$\bar{x}$: é a média da variável taxa de condomínio

$\bar{y}$: é a média da variável média de dias

O teste segue os seguintes passos:

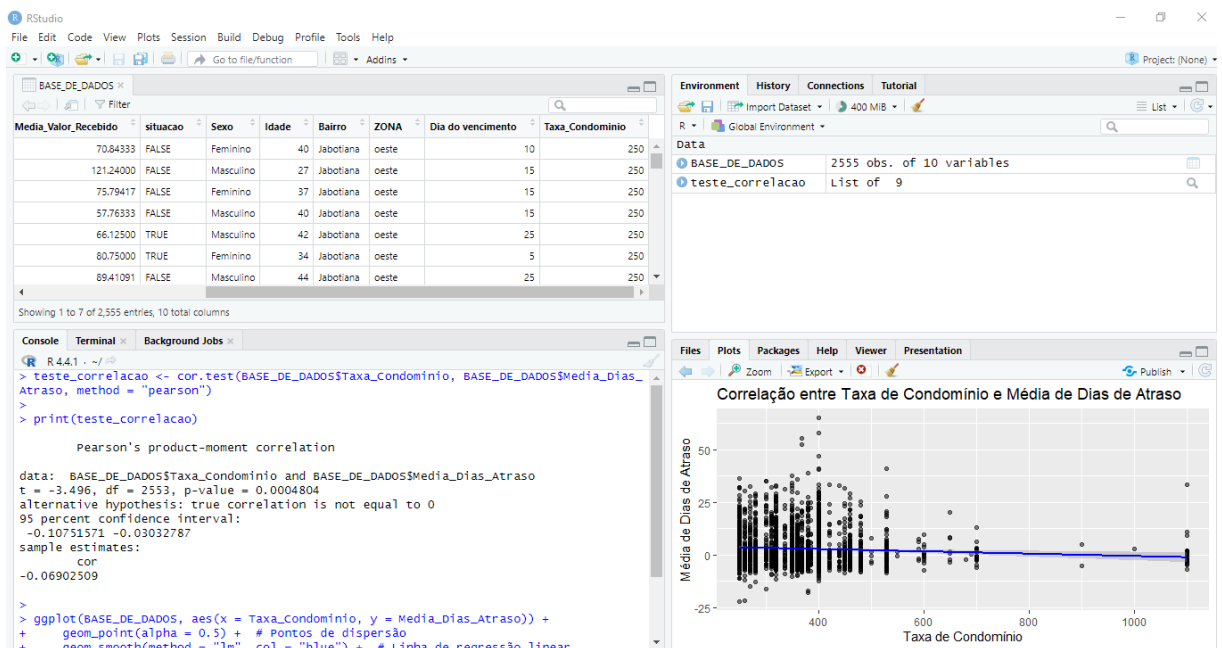
Figura 13- teste de correlação de Pearson para taxa de condomínio no R

A) importar a base de dados no software



Fonte: elaborado pelo autor

B) execução do teste no software



Fonte: elaborado pelo autor

4.3.3 Interpretação dos resultados fornecidos

Resultados do Teste:

- Estatística t: -3.4994
- Graus de liberdade (df): 2553
- Valor de r (correlação): -0.0691

- Valor-p: 0.0004742
- Intervalo de Confiança de 95% para r : -0.1076 a -0.0304

Interpretação

Valor de Correlação

o valor de $r = -0.0691$ indica que existe uma correlação negativa fraca (Devore, 2011). entre a taxa de condomínio e a média de dias de atraso/adiantamento. Essa correlação negativa sugere que, à medida que a taxa de condomínio aumenta, há uma leve tendência de redução no número de dias de atraso nos pagamentos. Em outras palavras, clientes que pagam taxas de condomínio mais altas tendem, de forma sutil, a quitar suas faturas com menor atraso ou até com um pouco mais de antecedência

Estatística t:

A estatística t de -3.4994 é usada para testar a significância da correlação. Esse valor negativo reflete a direção da correlação (negativa), mas seu valor em termos de magnitude indica que a correlação observada é estatisticamente significativa (Montgomery; Runger, 2014).

Valor-p

O valor-p é 0.0004742, que é muito menor que o nível de significância padrão de 0,05, concluindo que o teste foi estatisticamente significativo (Fisher, 1935). Isso significa que há evidências estatisticamente significativas para rejeitar a hipótese nula (H_0), que afirmava que não há correlação linear entre a taxa de condomínio e a média de dias de atraso/adiantamento.

Intervalo de Confiança de 95%

O intervalo de confiança para a correlação vai de -0.1076 a -0.0304. Como este intervalo não inclui 0, isso reforça ainda mais a conclusão de que a correlação observada é estatisticamente significativa. Além disso, como ambos os limites do

intervalo são negativos, isso confirma a direção da correlação negativa (Agresti, 2007).

4.3.4 Decisão

Com base nas interpretações rejeitamos a hipótese nula (H_0) e aceitamos a hipótese alternativa (H_1), concluindo que embora seja leve (-0.0691) há uma correlação estatisticamente significativa entre a taxa de condomínio e a média de dias de atraso/adiantamento.

4.4 COMPARAÇÃO DOS RESULTADOS COM OUTROS ESTUDOS

Nesta seção, os resultados obtidos neste estudo foram comparados com os achados de pesquisas anteriores, apresentados na seção 2.6 "Outros Estudos". Essa comparação tem como objetivo entender as semelhanças e diferenças entre as variáveis que influenciam a inadimplência nos diferentes contextos analisados.

4.4.1 Comparação com o Estudo de Baldi e Gianetti (2018)

Baldi e Gianetti (2018) estudaram os fatores que influenciam a inadimplência em uma cooperativa de crédito, e identificaram que as mulheres têm uma menor probabilidade de inadimplência do que os homens, além de concluírem que a escolaridade influencia no comportamento financeiro. No entanto, os resultados deste estudo não indicaram uma diferença significativa entre os gêneros dos clientes da empresa de internet quanto à inadimplência, sugerindo que, neste contexto, o gênero não é um fator determinante para o comportamento financeiro.

Embora Baldi e Gianetti(2018) tenham destacado a relevância da escolaridade, este estudo não analisou diretamente esse fator. Mesmo assim, variáveis como a idade e o valor das mensalidades podem fornecer indícios sobre a estabilidade financeira dos clientes, que pode estar relacionada ao nível de instrução.

4.4.2 Comparação com o Estudo de Higuchi (2023)

Higuchi (2023) conduziu uma análise em uma *fintech* e observou que clientes mais jovens (entre 18 e 25 anos) têm maior propensão à inadimplência, o que se alinha com os resultados deste estudo, onde essa mesma faixa etária apresentou os

maiores índices de inadimplência. Isso reforça a ideia de que, em ambos os contextos, a falta de experiência ou educação financeira pode estar contribuindo para a inadimplência entre os clientes mais jovens.

Higuchi (2023) também destacou que o uso de cartões de crédito está associado a um aumento na inadimplência. Embora este estudo não tenha investigado o impacto dos métodos de pagamento, essa variável pode ser relevante para estudos futuros no contexto da empresa de internet.

4.4.3 Comparação com o Estudo de Ferreira (2020)

Ferreira (2020) investigou a inadimplência entre profissionais de enfermagem e concluiu que indivíduos com maior renda, mas com alto comprometimento financeiro, tendem a ser mais inadimplentes. No presente estudo, verificou-se que clientes que pagam mensalidades mais altas também têm maior tendência à inadimplência, o que pode estar relacionado a um maior comprometimento financeiro, semelhante ao que foi observado por Ferreira (2020).

Embora a renda dos clientes não tenha sido diretamente analisada, o valor das mensalidades pode ser considerado um indicativo do nível de comprometimento financeiro, sugerindo a necessidade de medidas preventivas, como programas de educação financeira.

4.4.4 Comparação com o Estudo de Dynan, Skinner e Zeldes (2000)

Dynan, Skinner e Zeldes (2000) investigaram como despesas fixas, como taxas de condomínio, afetam a capacidade financeira das famílias. O estudo demonstrou que essas despesas consomem uma parte significativa da renda, limitando a flexibilidade para cumprir outras obrigações financeiras.

Os resultados do presente estudo apontam uma correlação leve entre o pagamento de taxas de condomínio e a inadimplência. Clientes que pagam taxas de condomínio mais altas tendem a manter seus pagamentos em dia, o que sugere que despesas fixas podem ser priorizadas, conforme sugerido por Dynan, Skinner e Zeldes (2000). No entanto, se essas despesas consomem uma parcela excessiva da renda, isso pode afetar a capacidade de honrar outras obrigações, como os serviços de internet.

4.4.5 Síntese Comparativa

De modo geral, os resultados do presente estudo estão em consonância com alguns dos achados dos estudos revisados, especialmente no que se refere à maior inadimplência entre clientes mais jovens, conforme identificado por Higuchi (2023), e a correlação entre despesas fixas e adimplência, como observado por Dynan, Skinner e Zeldes (2000). Entretanto, o impacto do gênero, destacado por Baldi e Gianetti (2018), não foi evidenciado neste estudo.

A análise comparativa ressalta a importância de considerar diferentes variáveis que afetam o comportamento financeiro dos clientes e sugere que a implementação de estratégias, como programas de educação financeira, pode ser uma solução eficaz para reduzir a inadimplência na empresa de internet.

5. CONCLUSÃO

Diante da análise realizada no presente estudo sobre o perfil de clientes inadimplentes em uma empresa de internet localizada em Aracaju/SE, é possível inferir conclusões valiosas que contribuem para uma compreensão mais abrangente dos padrões de pagamento e fatores associados. A pesquisa adotou uma abordagem quantitativa e descritiva, utilizando dados extraídos do Data Lake da empresa no período de janeiro a dezembro de 2022.

O estudo focou-se em diversas variáveis, como idade, gênero, data de vencimento, valor da mensalidade, localidade, e taxa de condomínio buscando identificar padrões que pudessem influenciar a adimplência ou inadimplência dos clientes. A metodologia incluiu a aplicação de testes de hipóteses, análise estatística descritiva e correlacional, proporcionando uma visão abrangente do comportamento dos clientes.

5.1 PRINCIPAIS RESULTADOS

A pesquisa revela informações importantes sobre o perfil dos clientes adimplentes e inadimplentes, destacando vários fatores que influenciam a inadimplência:

- **Valor das Mensalidades:** Observa-se uma correlação significativa entre o valor das mensalidades pagas e a situação de pagamento. Clientes que pagam valores mais altos demonstram uma tendência maior à inadimplência, evidenciada pela diferença estatisticamente significativa entre as médias dos valores pagos por clientes adimplentes (R\$ 82,57) e inadimplentes (R\$ 85,94)
- **Idade dos Clientes:** Verifica-se que existe uma correlação negativa, embora fraca, entre a idade dos clientes e a média de dias de atraso. Clientes mais jovens tendem a atrasar mais seus pagamentos, enquanto clientes mais velhos apresentam maior adimplência.
- **Gênero e Adimplência:** A análise não encontra evidências significativas de que o gênero dos clientes influencie a probabilidade de adimplência ou inadimplência.

- Taxa de Condomínio e Adimplência: Identifica-se uma correlação negativa fraca, mas estatisticamente significativa, entre a taxa de condomínio e a média de dias de atraso ou adiantamento dos pagamentos. Isso sugere que clientes com taxas de condomínio mais altas tendem a atrasar menos os pagamentos ou a pagarem com um pouco mais de antecedência, embora a relação seja leve. Esse comportamento pode estar associado a um perfil diferente de consumidor, no qual aqueles que residem em locais com taxas de condomínio mais elevadas possuam maior estabilidade financeira ou um planejamento de pagamento mais estruturado.
- Outro ponto que merece destaque é a observação indireta de que os clientes com mensalidades mais altas, que apresentaram maior tendência à inadimplência, não necessariamente coincidem com os clientes que residem em condomínios com taxas mais elevadas, os quais demonstraram maior propensão a pagar as faturas com antecedência. Esse contraste sugere que, possivelmente, os clientes que optam por planos mais caros de internet podem estar localizados em regiões ou condomínios com menor custo fixo, o que indica que o valor da mensalidade pode estar mais relacionado ao perfil de consumo do cliente do que ao seu poder aquisitivo.
- Adicionalmente, foi realizada uma análise de correlação separada por região para verificar se havia diferenças significativas entre as regiões. No entanto, os resultados não apresentaram correlações estatisticamente significativas $P > 0,05$, indicando que a variável “zona” não exerce uma influência diferenciada sobre a relação entre a taxa de condomínio e o comportamento de pagamento.

5.2 LIMITAÇÕES E SUGESTÕES PARA FUTURAS PESQUISAS

A pesquisa enfrenta limitações importantes, sendo a principal a falta de acesso a dados sobre a renda dos clientes. A ausência dessa informação essencial restringe a capacidade de explorar de forma mais aprofundada como diferentes níveis de renda afetam a inadimplência. Sem esses dados, torna-se impossível avaliar de forma abrangente o impacto econômico na capacidade de pagamento dos clientes.

Outras limitações identificadas incluem:

- Base de Dados Restrita: A pesquisa utiliza dados de uma única empresa, o que impossibilita a generalização dos resultados para outros contextos e setores.
- Métodos Quantitativos: A dependência de métodos quantitativos não permite captar todos os fatores que podem influenciar a inadimplência, como variáveis comportamentais e socioeconômicas específicas.

Para superar essas limitações, futuras pesquisas devem:

- Incluir Dados sobre a Renda dos Clientes: Coletar informações detalhadas sobre a renda dos clientes para permitir uma análise mais precisa da relação entre renda e inadimplência.
- Ampliar a Coleta de Dados: Expandir a amostra para incluir diferentes tipos de empresas e contextos regionais, garantindo uma maior representatividade e abrangência dos resultados.
- Utilizar Metodologias Diversificadas: Adotar abordagens qualitativas, como entrevistas e grupos focais, para captar fatores que os métodos quantitativos não conseguem explorar completamente, como motivações e comportamentos subjacentes.

5.3 CONSIDERAÇÕES FINAIS

Com base nos resultados obtidos nesta pesquisa, foi possível compreender melhor o perfil dos clientes inadimplentes e adimplentes de uma empresa fornecedora de internet em Aracaju/SE. Os testes estatísticos aplicados mostraram que variáveis como idade, valor da mensalidade e taxa de condomínio estão relacionadas à propensão à inadimplência.

Dentre os achados mais relevantes, destaca-se que clientes com mensalidades mais altas tendem a atrasar os pagamentos, enquanto aqueles que residem em condomínios com taxas condominiais mais elevadas tendem a manter suas contas em dia. Isso sugere que os fatores que influenciam a inadimplência vão além do valor cobrado pelo serviço, estando também ligados à gestão financeira pessoal e ao perfil econômico do local de residência.

Com base nesses resultados, propõem-se as seguintes ações estratégicas para a empresa:

1. Conceder descontos em mensalidades para clientes que residem em condomínios com taxas mais elevadas, uma vez que esse grupo demonstrou menor risco de inadimplência. Essa medida pode incentivar a fidelização e contribuir para um fluxo de caixa mais previsível.
2. Aplicar uma política de precificação diferenciada para clientes de condomínios com taxas mais modestas, especialmente os que contratam planos mais caros. Esses clientes revelaram um comportamento de consumo acima da sua capacidade de pagamento, resultando em atrasos. Nesse caso, a empresa pode considerar ajustar os valores desses planos ou oferecer alternativas mais compatíveis com o perfil financeiro da localidade.

Tais recomendações podem reduzir significativamente a inadimplência, tornando o modelo de negócios mais eficiente e alinhado ao perfil dos clientes. Os achados também oferecem base para políticas comerciais mais justas e estratégias de gestão mais eficazes, considerando o risco de cada público.

A pesquisa demonstra que a inadimplência pode ser prevenida com ações personalizadas e comunicação proativa, e reforça a importância da estatística como ferramenta de decisão estratégica. Apesar da ausência de dados sobre a renda dos clientes, os resultados servem como base para futuras pesquisas e ações gerenciais no setor.

Espera-se que as conclusões contribuam não apenas para a empresa analisada, mas também sirvam de referência para outras organizações com desafios semelhantes.

REFERÊNCIAS

ADMINISTRA BRASIL. **FLUXO DE CAIXA**. SÃO PAULO, p 15, 2018. Disponível em: <<https://administrabrasil.com.br/wp-content/uploads/2018/09/Fluxo-de-Caixa.pdf>>.

AGRESTI, A. **An introduction to categorical data analysis**. 2nd ed. Hoboken: Wiley, 2007

AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES. **Painéis de Dados de Outorga e Licenciamento**. Disponível em: <https://informacoes.anatel.gov.br/paineis/outorga-e-licenciamento>. Acesso em: 10 jul. 2023.

ALBUQUERQUE, P. H. M.; MEDINA, F. A. S.; SILVA, A. R. da. **Regressão Logística Geograficamente Ponderada Aplicada a Modelos de Credit Scoring**. Revista Contabilidade & Finanças, [S. l.], v. 28, n. 73, p. 93-112, 2017. DOI: 10.1590/1808-057x201703760. Disponível em: <https://www.revistas.usp.br/rcf/article/view/125577>. Acesso em: 10 set. 2023.

AMARAL, G. **Determinantes de inadimplência e de recuperação de crédito em um banco de desenvolvimento**. 2015. 307 f. Dissertação (Mestrado) - Curso de Administração, Universidade Federal de Minas Gerais, Belo Horizonte, 2015. Disponível em: https://repositorio.ufmg.br/bitstream/1843/BUBD-A9WECC/1/ppgadministracao_gustavohenriqueoliveiraamaral_dissertacaomestrado.pdf. Acesso em: 10 jul. 2023.

ANDRADE, S. F.; RIUL, P. H.; OLIVEIRA, M. S.; CAVALCANTI, M. F. **A inadimplência nas instituições particulares de ensino na cidade de Franca**. FACEF Pesquisa, v. 11, n. 1, 2008.

ANTONIAZZI, E. **Renda, endividamento e inadimplência: uma análise brasileira no período de 2010 a 2016**”, Revista OIDLES, 2019. n. Acesso em: 03 set 2023. Disponível em: <https://www.eumed.net/rev/oidles/27/renda-brasileira.html>

AZIZ.O. **Vade Mecum**.7. ed, Brasília,2023. Disponível em: <https://www2.senado.leg.br/bdsf/handle/id/608779>. Acesso em 13 ago.2023.

BAIL, C. **O live marketing como estratégia segmentada para o público jovem de curitiba**. 2014. 27 f. Monografia (Especialização) - Curso de Marketing Empresarial, Universidade Federal do Paraná, Curitiba, 2014. Disponível em: <https://acervodigital.ufpr.br/handle/1884/46621>. Acesso em: 10 set. 2023.

BALDI, M; GIANETTI, G. **Determinantes da inadimplência de pessoas físicas em cooperativa de crédito em Santa Catarina**. Quaestum, Santa Catarina, v. 2, n. 1, p. 1-13, 31 mar. 2018. Disponível em: <https://revista.ipecege.com/quaestum/article/view/586>. Acesso em: 24 out. 2023.

BARTH, N. **Inadimplência: construção de modelos de previsão**. São Paulo: Nobel, 2004. Disponível em: <https://books.google.com.br/books?id=XoDFLUiSvqC&lpq=PP1&hl=pt-BR&pg=PP1#v=onepage&q&f=false>. Acesso em 13 ago. 2023.

BAUMAN, Z. **Vida para consumo: a transformação das pessoas em mercadoria**. Rio de Janeiro: Jorge Zahar Editora, 2008.

CAETANO, A et al. **Consumismo e endividamento**. 7ª Jornada Acadêmica – Crescimento Regional: Inovação e Tecnologia no Mercado de Trabalho, Santa Helena, v. 7, n. 1, p. 1-3, 11 abr. 2016. Disponível em: <https://anais.ueg.br/index.php/jaueg/article/view/6691>. Acesso em: 03 set. 2023.

CARVALHO, t. **Sistema Capitalista**. Politize. 2019. Disponível em: <https://www.politize.com.br/sistema-capitalista-origem/> Acesso em 03 de setembro de 2023

CLAUDINO, L.. **Educação financeira e endividamento**: um estudo de caso com servidores de uma instituição pública. Anais do Congresso Brasileiro de Custos - ABC, 2009[I.], Disponível em: <https://anaiscbc.emnuvens.com.br/anais/article/view/1029>. Acesso em: 19 nov. 2023

CONFEDERAÇÃO NACIONAL DE COMÉRCIO DE BENS, SERVIÇOS E TURISMO, **Pesquisa de Endividamento e Inadimplência do Consumidor**, 2023. Disponível em: <https://pesquisascnc.com.br/pesquisa-peic/>. Acesso em: 20 jul. 2023.

DEVORE, J. L. **Probability and Statistics for Engineering and the Sciences**. 8th ed. Boston: Cengage Learning, 2011

DYNAN, K. E.; SKINNER, J.; ZELDES, S. P. **Do the rich save more?** Relatório de Pesquisa. National Bureau of Economic Research, Cambridge, MA, v. 1, n. 1, p. 1-61, 10 set. 2000. Disponível em: https://www.researchgate.net/publication/24104353_Do_the_Rich_Save_More. Acesso em: 31 mar. 2024.. Disponível em: https://www.researchgate.net/publication/24104353_Do_the_Rich_Save_More/link/0deec525c78c7e84c7000000/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19. Acesso em: 31 mar. 2024.

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R. **Advances in Knowledge Discovery & Data Mining** 1 ed. American Association for Artificial Intelligence, Menlo Park, Califórnia, 1996. 611 folhas. Disponível em: https://www.researchgate.net/publication/2460134_KDD_for_Science_Data_Analysis_Issues_and_Examples. Acesso em: 12 set. 2023.

FERREIRA, C. **Estudo sobre as causas da inadimplência dos profissionais de enfermagem inscritos no coren/am**. 2020. 192 f. Dissertação (Mestrado) - Curso de Economia, Universidade de Brasília, Brasília, 2020. Disponível em: <https://repositorio.unb.br/handle/10482/39671?locale=es>. Acesso em: 20 out. 2023.

FERREIRA, J. **A importância da educação financeira pessoal para a qualidade de vida**. Caderno de Administração, Bauru, v. 11, n. 1, p. 1-17, 01 dez. 2017. Disponível em: <https://revistas.pucsp.br/caadm/article/view/33268>. Acesso em: 04 set. 2023.

FISHER, R. A. **The Design of Experiments**. Edinburgh: Oliver and Boyd, 1935.

GANSKE, L. **O perfil de clientes inadimplentes**: estudo de caso em microempresa do ramo de calçados e confecções. 2020. 31 f. Monografia (Especialização) - Curso

de Gestão Contábil e Financeira, Universidade Tecnológica Federal do Paraná, Pato Branco, 2020. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/26233>. Acesso em: 11 set. 2023.

GIL, A. **Como elaborar projetos de pesquisa**. 4ª. ed. São Paulo: Atlas S/A, 2002

GIL, A. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo, SP: Atlas, 2010

GOMES, Carlos Roberto. **GESTÃO DE CRÉDITO E COBRANÇA. Introdução à gestão de riscos: ênfase em cadastro, crédito e cobrança**. São Paulo: Editora Viena, 2013. 160 p

GUTERRES, Everton Luís de Lima. **Inadimplência, um estudo de caso: o delineamento do perfil**. 2009. 12 f. TCC (Graduação) - Curso de Administração, Faculdade Integrada de Santa Maria, Santa Maria, 2009. Disponível em: <https://saeopro.ufv.br/wp-content/uploads/2015/06/2010.5.pdf>. Acesso em: 12 set. 2023.

HAN, J.; KAMBER, M. **Data mining: concepts and techniques**. 2. edição. Morgan Kaufmann, 2006.

HAN, J.; KAMBER, M.; JIAN, P. **Data mining concepts and techniques**. 3. ed. Waltham: Elsevier, 2012.

HAN, J.; PEI, J.; KAMBER, M. **Data mining: concepts and techniques**. 4th ed. Elsevier, 2022.

HIGUCHI, C. **Estudo da correlação das características de clientes de empresa de meio de pagamento com a inadimplência**. 2023. 64 f. Dissertação (Doutorado) - Curso de Administração, Fundação Getulio Vargas, São Paulo, 2023. Disponível em: <https://bibliotecadigital.fgv.br/dspace/handle/10438/33948>. Acesso em: 25 out. 2023.

LEHMANN, E. L.; romano, j. P. **Testing Statistical Hypotheses**. 3rd ed. New York: Springer, 2005.

MATIAS, A. B.; DAUBERMANN, E. C.; RICCI, E. T. **Inadimplência não é insolvência: um estudo conceitual e empírico com uma abordagem macrofinanceira**. Revista FACEF Pesquisa, v. 12, n. 1, p. 36-51, 2009. Disponível em: <http://periodicos.unifacef.com.br/index.php/facefpesquisa/article/viewFile/4/200>. Acesso em: 17 ago. 2023.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied statistics and probability for engineers**. 6. ed. Hoboken: Wiley, 2014.

MOORE, D. S.; McCABE, G. P.; CRAIG, B. A. **Introduction to the Practice of Statistics**. 7th ed. New York: W.H. Freeman and Company, 2012.

Mian, A., Rao, K., & Sufi, A. (2013). **Household balance sheets, consumption, and the economic slump**. Quarterly Journal of Economics, 128(4), 1687-1726.

PEARSON, K. **Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia**. Philosophical Transactions of the Royal

Society of London. Series A, Containing Papers of a Mathematical or Physical Character, v. 187, p. 253-318, 1896.

PINTO, N. G. M.; CORONEL, D. A. **Inadimplência no serviço de TV por assinatura: um estudo sobre as suas causas em uma empresa do setor da cidade de Santa Maria** - RS. SINERGIA - Revista do Instituto de Ciências Econômicas, Administrativas e Contábeis, [S. l.], v. 16, n. 2, p. 41–51, 2013. Disponível em: <https://periodicos.furg.br/sinergia/article/view/3671>. Acesso em: 29 jul. 2023.

REIS.A. **Os principais fatores que influenciam no endividamento e inadimplência das famílias brasileiras**. 2020. 20 f. TCC (Graduação) - Curso de Administração, Centro Universitário Fametro, Fortaleza, 2020. Disponível em: <http://repositorio.fametro.com.br/jspui/handle/123456789/164>. Acesso em: 02 set. 2023.

ROCHA, E. **Segmentação do Perfil de Clientes Inadimplentes Utilizando Ferramentas Computacionais**. 2021. 79 f.- Curso de Estatística, Universidade Estadual Paulista, São Paulo, 2022. Disponível em: <https://repositorio.unesp.br/handle/11449/217472>. Acesso em: 11 ago. 2023.

RUBERTO, I. et al. **A influência dos fatores macroeconômicos sobre o endividamento das famílias brasileiras no período 2005-2012**. Estudos do CEPE, p. 58-77, 28 ago. 2013. Acesso em 04 set 2023. Disponível em: <https://online.unisc.br/seer/index.php/cepe/article/view/3328>.

SANDRONI, P. **Novíssimo Dicionário de Economia**. São Paulo: Best Seller, 1999

SERASA. **Mapa da Inadimplência e Negociação de Dívidas no Brasil**. Brasil: Serasa, 2022. Color. Disponível em: <https://www.serasa.com.br/limpa-nome-online/blog/mapa-da-inadimplencia-e-renogociacao-de-dividas-no-brasil/>. Acesso em: 10 set. 2023.

SILVA, A et al. **Qualidade de vida e endividamento**. Desafio Online, Campo Grande, v. 8, n. 2, p. 353-377, 20 jul. 2020. Disponível em: <https://desafioonline.ufms.br/index.php/deson/issue/view/599>. Acesso em: 03 set. 2023.

SILVA, A. L.; COSTA, R. M.; OLIVEIRA, T. P. **Advanced machine learning applications in KDD processes**. Journal of Data Science, v. 21, n. 1, p. 45-60, 2023.

SPC BRASIL. **Perfil do inadimplente e das dívidas no brasil**. Brasil: Spc, 2016. Color. Disponível em: <https://www.spcbrasil.org.br/wpimprensa/wp-content/uploads/2016/08/An%C3%A1lise-Perfil-do-Inadimplente-e-da-D%C3%ADvida.pdf>. Acesso em: 10 set. 2023.

TAVARES, S. **A inadimplência: como administrar**. Revista de Educação do Cogeime, Piracicaba, v. 5, n. 9, p. 105-110, 28 mar. 1996. DISPONÍVEL EM: <HTTPS://WWW.METODISTA.BR/REVISTAS/REVISTAS-COGEIME/INDEX.PHP/COGEIME/ISSUE/VIEW/28/SHOWTOC>. ACESSO EM: 20 AGO. 2023.

TOCHETTO, E. **A influência da localização geográfica na inadimplência das empresas atendidas por um banco de varejo utilizando a regressão logística**

geograficamente ponderada. 2018. 94 f. Dissertação (Mestrado) - Curso de Mestrado em Administração, Universidade Positivo, Curitiba, 2018. Disponível em: <https://repositorio.unifran.edu.br/jspui/handle/123456789/1919>. Acesso em: 10 set. 2023.

VIANA, E. J.; SILVA, J. T.; SANTOS, M. F. M. **Estudo sobre a inadimplência na empresa Ampernet: telecomunicações no município de Ampére-PR.** Trabalho de Conclusão de Curso (Graduação) – Curso de Administração, Faculdade de Ampére, Ampére, 2019. 23 f. Disponível em: <http://www.famper.com.br/tcc-online/submissao/49>. Acesso em: 11 set. 2023.

VIEIRA, A et al. **O custo do crédito pessoal em relação ao nível de endividamento das famílias brasileiras e à taxa de juros selic** doi. Revista Economia & Gestão, Belo Horizonte, v. 14, n. 36, p. 36-138, 23 out. 2014. Acesso em: 03 set. 2023. Disponível em: <http://periodicos.pucminas.br/index.php/economiaegestao/article/view/P.1984-6606.2014v14n36p138>. Acesso em: 03 set. 2023.

WELCH, B. L. **The generalization of "Student's" problem when several different population variances are involved.** *Biometrika*, v. 34, n. 1/2, p. 28-35, 1947.

WITTEN, I. H.; FRANK, E.; HALL, A. M. **Data mining: practical machine learning tools and techniques.** 3. ed. Burlington: Elsevier, 2011. Disponível em: http://academia.dk/BiologiskAntropologi/Epidemiologi/DataMining/Witten_and_Frank_DataMining_Weka_2nd_Ed_2005.pdf. Acesso em: 14 set. 2023.

ZERRENNER, S. **Estudo sobre as razões para o endividamento da população de baixa renda.** 2007. Dissertação (Mestrado em Administração) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2007. doi:10.11606/D.12.2007.tde-13112007-120236.

