

Jordão Flores Finamor

**Pesquisa Exploratória sobre Operações de
Crédito no Brasil: Uma Perspectiva de 2019 a
2022.**

Brasil

27 de Julho de 2023

Jordão Flores Finamor

Pesquisa Exploratória sobre Operações de Crédito no Brasil: Uma Perspectiva de 2019 a 2022.

Resumo Expandido em conformidade com as
normas ABNT

Universidade Federal do Rio Grande do Sul - UFRGS

Grupo de Pesquisa em Analytics

Curso de Especialização em Business Analytics

Brasil

27 de Julho de 2023

Sumário

1	INTRODUÇÃO	3
	Introdução	3
1.1	Objetivos	4
2	METODOLOGIA	5
3	RESULTADOS E DISCUSSÃO	6
4	CONCLUSÃO	13
	REFERÊNCIAS	14
	ANEXOS	15
	ANEXO A – CÓDIGO UTILIZADO	16

1 Introdução

Olhando em retrospectiva, a partir do ano de 2023, o mercado financeiro nacional passou por períodos de turbulência e muito receio por parte de analistas financeiros e investidores durante os últimos quatro anos. Segundo comunicado do governo brasileiro em [Portal Gov.br \(2020\)](#), a Organização Mundial da Saúde classifica coronavírus como pandemia. O mercado financeiro mundial foi diretamente afetado pelo avanço dos casos de coronavírus e a crise política causada. Além disso, o mercado interno também foi impactado por outros problemas, como em 2020, tivemos uma das maiores quedas do PIB na vida brasileira (-4,3%), com contração de 11% no segundo trimestre e a maior crise desde a década de 1930, explica [Solutis \(2021\)](#). Já em 2021 e 2022, conforme [Praisce \(2022\)](#), o país passou por uma crise político-econômica afetada por diversos fatores, tais como: crise fiscal interna, reajuste do funcionalismo, aumento dos juros, quebra do teto de gastos, impactos da variante Ômicron do vírus COVID-19 e conflito armado entre Rússia e Ucrânia.

Este estudo teve como objetivo realizar uma pesquisa exploratória utilizando os dados agregados de operações de crédito divulgados mensalmente pelo Banco Central do Brasil através do Sistema de Informações de Créditos (SCR). O propósito foi observar o comportamento das operações de crédito no Brasil, a pesquisa envolveu a aplicação de técnicas de estatística descritiva e testes de hipóteses para validar e analisar os dados ao longo do tempo.

Por meio da estatística descritiva, foram resumidas e descritas as principais características dos dados, incluindo média, mediana, desvio padrão. Além disso, foram realizados testes de hipóteses para avaliar a significância estatística de determinadas suposições e diferenças entre grupos e períodos.

A análise realizada permite obter perspectivas valiosas sobre o comportamento das operações de crédito ao longo do tempo, considerando variáveis como tipo de cliente (Pessoa Física/Pessoa Jurídica), modalidade de crédito, estado, Classificação Nacional de Atividades Econômicas (CNAE) para pessoas jurídicas, natureza da ocupação para pessoas físicas, porte/rendimento dos clientes, origem de recursos e indexador das operações.

É importante ressaltar que este estudo adere aos pressupostos dos testes estatísticos aplicados e explicados na Metodologia. Todas as análises foram devidamente documentadas para garantir a transparência e replicabilidade dos resultados.

1.1 Objetivos

O objetivo deste trabalho é analisar as operações de crédito do sistema financeiro do Brasil nos últimos quatro anos, no período de 2019 a 2022, com o propósito de quantificar e compreender o comportamento dessas operações.

2 Metodologia

Neste estudo, adotou-se uma abordagem metodológica que envolveu a coleta de dados das operações de crédito do sistema financeiro do Brasil entre o período de 2019 a 2022, utilizando o Sistema de Informações de Créditos (SCR) do Banco Central do Brasil, disponível em seu site de dados abertos [Banco Central do Brasil \(2023\)](#). A análise desses dados permitirá um entendimento mais completo sobre o comportamento dessas operações de crédito durante o período de pandemia global e também de crises econômicas no mercado nacional.

Esses dados baseiam-se no layout do documento 3040 do Sistema de Informações de Créditos (SCR), com informações detalhadas de todas as operações de crédito cursadas no país com valor superior a R\$ 200,00. Para realizar a exploração, processamento, modelagem e análise dos dados, utilizou-se a linguagem de programação em *R*, que, segundo [Gardener \(2012\)](#) é uma linguagem de programação estatística gratuita, com código fonte livre, poderosa e flexível.

Na parte de análise de dados, foram aplicadas técnicas de estatística descritiva. Conforme [Santos \(2007\)](#) relata, a estatística descritiva busca sintetizar valores de mesma natureza para obter uma visão global sobre a sua variação. Além disso, foram utilizados técnicas de separação dos dados em períodos diferentes, para observar a diferença entre eles e formular testes de hipóteses que possam responder à diferença de comportamento entre os anos. Para investigar a presença de diferenças significativas no comportamento das operações de crédito ao longo dos períodos em estudo, empregaram-se técnicas estatísticas como o teste de Dunn e o teste de Kruskal-Wallis.

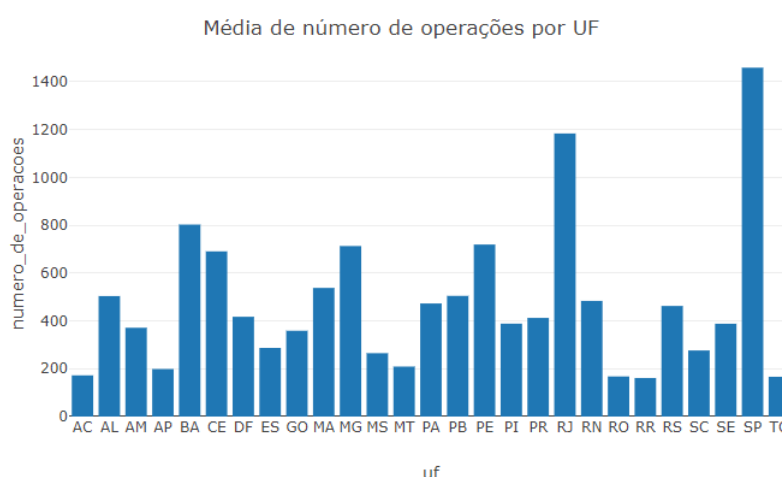
[Conover \(1999\)](#) explica que o teste de Dunn é uma extensão do teste de Wilcoxon-Mann-Whitney, utilizado para comparação múltipla de grupos independentes. Ele permite identificar pares de grupos que apresentam diferenças estatisticamente significativas entre si, mesmo quando há mais de dois grupos envolvidos. Além disso, o autor também explica que o teste de Kruskal-Wallis é uma extensão não paramétrica do teste de análise de variância ANOVA (*Analysis of Variance*) ou análise de variância. Kruskal-Wallis é utilizado para comparar três ou mais grupos independentes. Ele avalia se há diferenças estatisticamente significativas entre as medianas dos grupos, sendo uma alternativa robusta quando os dados não atendem aos pressupostos de normalidade e homogeneidade de variâncias exigidos pela ANOVA.

3 Resultados e Discussão

Levando em consideração que a base de dados possui um grande volume de dados a serem analisados e carregados, foi primeiramente realizado uma análise do primeiro mês e do primeiro ano dentro do período estipulado entre 2019 e 2023, janeiro de 2019. Nesta amostra inicial, já foi possível identificar um enorme volume de dados: 623.580 linhas de dados distribuídas entre 23 colunas com grupos.

Dando início ao trabalho de exploração, foram gerados gráficos sobre essa amostra, visando entender o comportamento geral dos dados. A coluna operações representa o número de operações de crédito para uma dada série. Este dado é relevante para compararmos com as categorias presente na base de dados utilizada neste estudo. Na [Figura 1](#) é exposto uma descrição do volume de operações por estado.

Figura 1 – Volume de operações de crédito por estado.

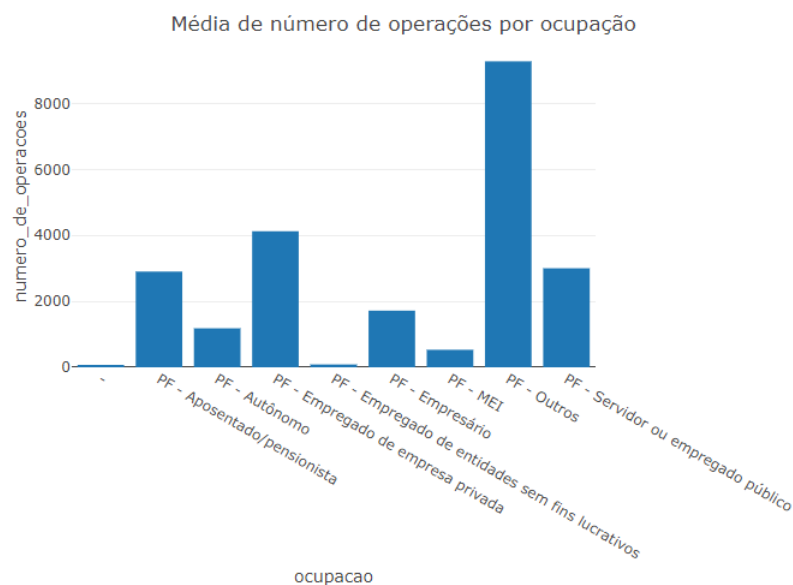


Fonte: autor.

Sobre o volume das operações de crédito separadas por natureza da ocupação de pessoas físicas, o valor de maior relevância foi a categoria outros. Esta categoria consiste em ocupações que não entram nos demais conjuntos apresentados na [Figura 2](#). Na sequência temos os grupos de PF - empregado de empresa privada, PF - Servidor ou empregado público e PF - Aposentado/pensionista.

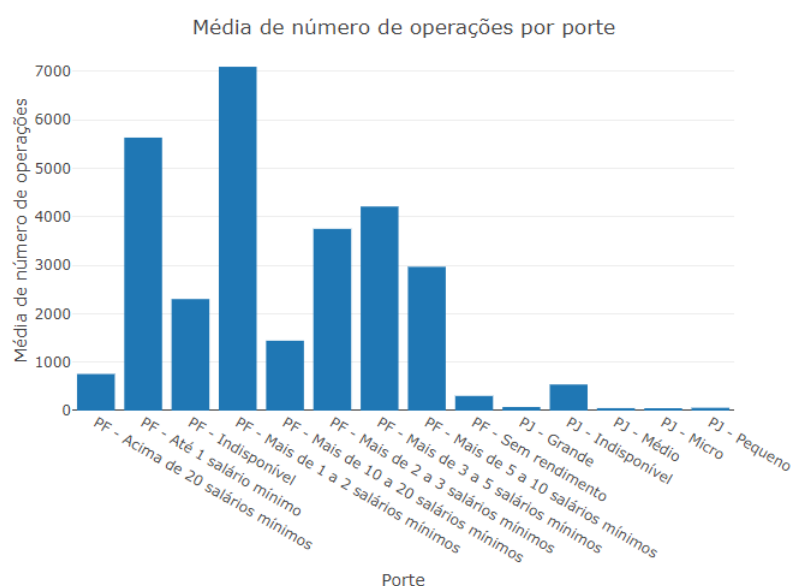
Ainda sobre pessoas físicas, analisamos este perfil separado por porte, conforme [Figura 3](#). Nota-se que as barras com maiores médias de número de operações são as de porte Pessoa Física, com números muito mais expressivos que o porte Pessoa Jurídica. O maior porte foi o de 1 a 2 salários mínimos, e em segundo lugar, até 1 salário mínimo, o que indica onde está a concentração de operações de cartão de crédito do Brasil.

Figura 2 – Volume de operações de crédito por natureza da ocupação de pessoas físicas.



Fonte: autor.

Figura 3 – Volume de operações de crédito por porte de pessoas físicas.

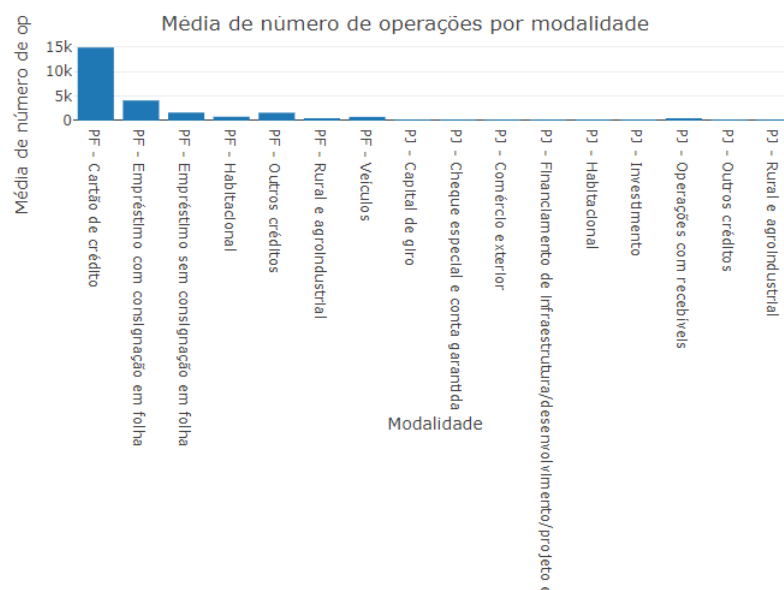


Fonte: autor.

Sobre a modalidade das operações de crédito, observamos na [Figura 4](#) que os dados trazem a média de número de operações para pessoas físicas e jurídicas. Neste gráfico, observamos um volume muito expressivo na modalidade pessoa física - cartão de crédito, e a seguir PF - Empréstimo com consignação em folha.

Essas observações iniciais foram importantes para identificar o comportamento

Figura 4 – Volume de operações de crédito por modalidade.



Fonte: autor.

geral dos dados a serem estudados de maneira aprofundada. Os pontos principais escolhidos como foco de estudo foram número de operações de créditos por estado, e também os dados na modalidade cartão de crédito (pessoa física). Primeiro, para observar se existe uma uniformidade no comportamento das operações de crédito entre os estados, e em um segundo momento, para observar o comportamento da modalidade que possui, com uma boa margem de diferença, o maior volume de operações de crédito.

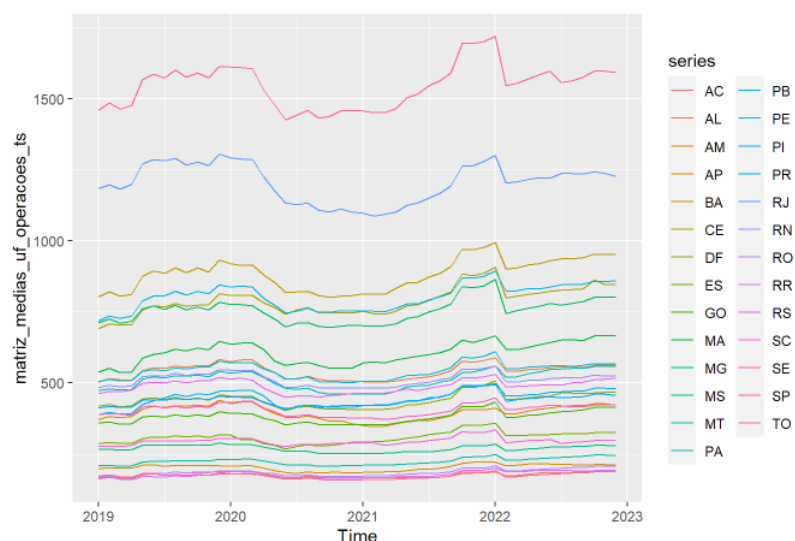
Considerando o volume massivo de dados a serem lidos, foi necessário fazer a leitura de cada planilha com os dados mensais, calcular e armazenar o valor médio do volume de operações. Esta média foi calculada por estado e armazenada em uma matriz. Posteriormente, foi alterado a matriz que contém esses dados com as médias em uma série temporal, para observar o comportamento do período. Esta série temporal está representada graficamente na [Figura 5](#).

Esta mesma lógica foi usada uma segunda vez, para computar e analisar os dados de operações de crédito sobre a ótica da modalidade cartão de crédito (pessoa física), conforme [Figura 6](#).

Observa-se em ambas séries temporais, [Figura 5](#) e [Figura 6](#), uma queda significativa no volume de operações de crédito nos períodos de 2020 e 2022. Existem algumas hipóteses que podem ser consideradas na busca por uma resposta sobre estas quedas. Conforme o site Solutis:

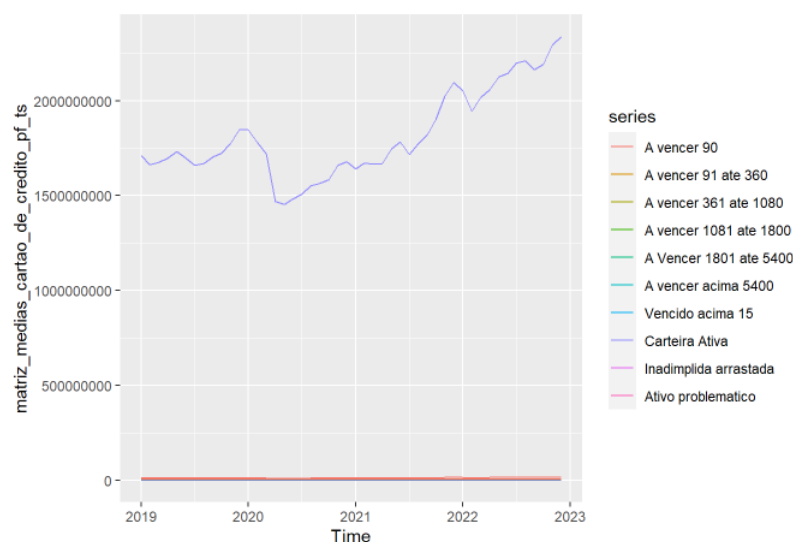
“O ano que não terminou para o setor de serviços financeiros”. Essa, certamente, será a maneira pela qual o ano de 2020 será lembrado, em virtude dos impactos da pandemia do novo coronavírus no mercado

Figura 5 – Série temporal com a média de operações de crédito por estado dos últimos 4 anos.



Fonte: autor.

Figura 6 – Série temporal com a média de operações de crédito com modalidade cartão de crédito (pessoa física) dos últimos quatro anos.



Fonte: autor.

financeiro. [...] Em 2020 tivemos uma das maiores quedas do PIB na vida brasileira (-4,3%), e igualamo-nos aos exorcizados anos da década de 90 (-4,35%) e 1981 (-4,25%). Vale registrar, ainda, a impressionante contração de 11% no segundo trimestre. Em suma, entidades multilaterais, nações e economistas reputam que estamos enfrentando a maior crise desde a década de 1930 — e mais severa que a de 2009, quando estourou a bolha do mercado imobiliário. [...] Ao final de 2020, diversas vacinas estavam em teste pelo mundo. E 2021 começaria com altas expectativas de acelerar

o fim da pandemia, além da recuperação da economia e dos mercados. A realidade, porém, foi um pouco diferente. As variantes do vírus, uma segunda onda forte e o atraso na vacinação, além da piora do ambiente econômico e fiscal, devem criar obstáculos. (Solutis, 2021)

Sobre o período de 2022, observou-se uma queda significativa em ambos os casos da Figura 5 e Figura 6. O site Praisce traz uma perspectiva bem interessante também sobre a situação do mercado financeiro neste período em que identificamos queda:

O ano de 2022 começou bem agitado no mercado financeiro mundial, tendo como principais fatos a explosão de casos de COVID-19 no mundo inteiro por conta da variante Ômicron, a perspectiva de alta de juros do FED (Banco Central dos Estados Unidos) e o potencial conflito armado entre Rússia e Ucrânia. Já no noticiário doméstico, a situação continua seguindo a mesma linha do que vinha ocorrendo em 2021: crise fiscal, reajuste do funcionalismo, aumento dos juros, quebra do teto de gastos, eleições, inflação nas alturas e para coroar tudo isso, número recorde de casos de COVID-19. (Praisce, 2022)

Para observarmos se essas ou quaisquer outras hipóteses possuem relevância, primeiro precisamos aprofundar a análise sobre os dados. Para simplificarmos, foi escolhido o estado do Rio Grande do sul, considerando que todos os estados possuem um comportamento semelhante. Dividiu-se os dados entre três períodos dentro do intervalo de 2019 a 2023, conforme Tabela 1.

Tabela 1 – Teste de Kruskal-Wallis sobre operações de crédito do Rio Grande do Sul.

Qui-quadrado	Grau de Liberdade	P-valor
26.129	2	0.000002119

Fonte: autor.

Com base no resultado do teste de Kruskal-Wallis, o valor do teste de qui-quadrado (chi-squared) é igual a 26.129. O qui-quadrado é calculado a partir das diferenças entre os valores observados e esperados sob a hipótese nula. Quanto maior for o valor do qui-quadrado, maior será a evidência de que há diferenças estatisticamente significativas entre as medianas dos grupos. O grau de liberdade (df) é igual a 2. Quanto maior o grau de liberdade, mais "flexível" é a distribuição qui-quadrado, ou seja, ela se torna mais assimétrica e se aproxima de uma distribuição normal.

O P-valor é igual a 0.000002119, o que indica que o valor-p é menor que o nível de significância comum de 0.05, concluindo que há evidências estatísticas para rejeitar a hipótese nula de que as distribuições da variável RS são iguais nos três períodos. Em outras palavras, existe uma diferença significativa na distribuição da variável de operações de crédito do estado do Rio Grande do Sul entre os 3 períodos.

Para analisar esta diferença de distribuição, vamos comparar os períodos entre si através do teste de Dunn, conforme Tabela 2 e Tabela 3.

Tabela 2 – Estatística de Dunn sobre operações de crédito nos 3 períodos do Rio Grande do Sul.

Período	1	2
2	3.901714	
3	-0.909137	-4.810851

Fonte: autor.

Tabela 3 – Valor do p-ajustado sobre operações de crédito nos 3 períodos do Rio Grande do Sul.

Período	1	2
2	0.0001*	
3	0.1816	0.0000*

Fonte: autor.

Ao compararmos o período 1 com o 2 observamos a estatística de Dunn em 3.901714 e o valor do p-ajustado em 0.0001*. Isso indica que há diferença estatisticamente significativa entre os períodos. Entre o período 1 com o 3 observamos a estatística de Dunn em -0.909137 e o valor do p-ajustado em 0.1816, ou seja, não há diferença estatisticamente significativa entre os períodos. Quando analisamos o período 2 com o 3, observamos a estatística de Dunn em -4.810851 e o valor do p-ajustado em 0.0000*, concluindo-se que há diferença estatisticamente significativa entre os períodos.

Seguindo a mesma lógica de separação em três períodos distintos para significância estatística, foi feita uma análise das operações de crédito por modalidade cartão de crédito (pessoa física), mostrado na [Tabela 4](#).

Tabela 4 – Teste de Kruskal-Wallis sobre operações de crédito na modalidade Cartão de Crédito - Pessoa Física

Qui-quadrado	Grau de Liberdade	P-valor
33.004	2	0.0000000681

Fonte: autor.

Sobre estas respostas, tiram-se algumas conclusões: O valor do teste de qui-quadrado de Kruskal-Wallis é igual a 33.004, com 2 graus de liberdade. O valor-p ajustado associado ao teste é extremamente baixo, com valor igual a 0.0000000681. Isso indica uma forte evidência estatística para rejeitar a hipótese nula de que não há diferenças significativas entre as distribuições dos períodos. Com base nesses resultados, podemos concluir que existe uma diferença significativa nas operações de cartão de crédito pessoa física entre os períodos analisados.

Utilizando-se do teste de Dunn para comparar os períodos entre si, obtemos os seguintes valores apresentados na [Tabela 5](#) e [Tabela 6](#).

A comparação entre período 1 e 2 apresenta uma estatística de Dunn de 1.464721 e um valor-p ajustado de 0.0715. Nesta observação conclui-se que não há diferença estatisticamente significativa. Sobre o período 1 e 3, a estatística de Dunn é de -4.078490 e o valor-p ajustado de 0.0000*, indicando uma diferença estatística significativa. Por fim, entre o período 2 e 3, a estatística de Dunn apresentou um valor de -5.543212 com o

Tabela 5 – Estatística de Dunn sobre operações de crédito na modalidade Cartão de Crédito - Pessoa Física.

Período	1	2
2	1.464721	
3	-4.078490	-5.543212

Fonte: autor.

Tabela 6 – Valor do p-ajustado sobre operações de crédito na modalidade Cartão de Crédito - Pessoa Física.

Período	1	2
2	0.0715	
3	0.0000*	0.0000*

Fonte: autor.

valor-p ajustado de 0.0000*, na qual se conclui que há diferença estatística significativa entre os períodos.

Além de ser possível observar essas quedas graficamente, e de aprofundar as análises de maneira a observar ela em três períodos distintos e se estas informações possuem relevância estatística, foram computadas as médias e desvio padrão dos períodos, conforme [Tabela 7](#) e [Tabela 8](#).

Tabela 7 – Média e mediana do volume de operações de crédito no Rio Grande do Sul.

Período	Média	Mediana
1	496.4444	18.26579
2	465.0103	11.24271
3	504.5259	14.22434

Fonte: autor.

Tabela 8 – Média e mediana do volume de operações de crédito na modalidade cartão de crédito (pessoa física).

Período	Média	Mediana
1	1.710.331.780	87.102.230
2	1.634.035.888	101.472.785
3	2.098.658.253	139.009.949

Fonte: autor.

Olhando sobre a ótica de operações de crédito na modalidade cartão de crédito pessoa física, e também sobre o estado do Rio Grande do Sul, observa-se que as operações de crédito no Brasil apresentaram uma mudança comportamental entre 2019 e 2023. Através da estatística descritiva e de testes de hipóteses, obtive fortes evidências estatísticas para concluir que houve variação significativa entre estes períodos ao compararmos de maneira segmentada entre três períodos.

4 Conclusão

Os resultados deste estudo representam um passo significativo na compreensão do comportamento das operações de crédito no Brasil ao longo do período de 2019 a 2023, com destaque para as mudanças observadas nos anos iniciais de 2020 e 2022. Ao utilizar testes estatísticos adequados, como o de Dunn e Kruskal-Wallis, pudemos identificar diferenças estatisticamente significativas entre os períodos analisados. Essas descobertas têm implicações importantes para o setor financeiro brasileiro e para a formulação de políticas públicas que impactam a economia em sua totalidade.

A análise estatística das operações de crédito fornece informações essenciais para os agentes econômicos, como bancos, investidores e empreendedores, que podem se beneficiar ao compreender as tendências e mudanças no cenário de crédito. Os resultados podem ajudar as instituições financeiras a ajustar suas estratégias e tomar decisões mais acuradas e assertivas ao conceder empréstimos e financiamentos.

Além disso, as informações obtidas podem contribuir para o desenvolvimento de políticas públicas mais efetivas no contexto econômico do Brasil. O governo e as autoridades financeiras podem utilizar esses dados para avaliar o impacto de medidas adotadas ao longo dos anos, como políticas de estímulo ao crédito ou de contenção econômica durante períodos de crise. Compreender como essas ações afetam as operações de crédito pode ser crucial para impulsionar o crescimento sustentável e promover a estabilidade econômica.

No entanto, é essencial ressaltar que a análise dos dados de operações de crédito é complexa e multidimensional. Para uma compreensão completa dos fatores que influenciaram as mudanças nas operações de crédito, é necessário explorar mais profundamente variáveis específicas que possam ter desempenhado um papel significativo. Fatores econômicos internos e externos, como taxas de juros, política fiscal, variações cambiais e eventos globais, como a pandemia de COVID-19, podem ter tido impacto direto nas decisões de crédito dos agentes econômicos.

Recomenda-se, portanto, que pesquisas futuras se aprofundem na análise de tais variáveis, utilizando modelos econométricos e técnicas estatísticas mais avançadas. Além disso, a expansão da base de dados para incluir informações adicionais, como segmentação de setores da economia ou tipos específicos de crédito, pode enriquecer ainda mais a compreensão dos fenômenos relacionados ao crédito no Brasil.

Referências

Banco Central do Brasil. *SCR Data*. 2023. Disponível em: <https://dadosabertos.bcb.gov.br/dataset/scr_data>. Citado na página 5.

CONOVER, W. J. *Practical nonparametric statistics*. [S.l.]: john wiley & sons, 1999. v. 350. Citado na página 5.

GARDENER, M. *Beginning R: the statistical programming language*. [S.l.]: John Wiley & Sons, 2012. Citado na página 5.

Portal Gov.br. *OMS classifica coronavírus como pandemia*. 2020. Disponível em: <<https://www.gov.br/pt-br/noticias/saude-e-vigilancia-sanitaria/2020/03/oms-classifica-coronavirus-como-pandemia>>. Citado na página 3.

Praisce. *Acontecimentos que movimentaram o mercado financeiro: janeiro de 2022*. 2022. Disponível em: <<https://www.praisce.com.br/acometimentos-que-movimentaram-o-mercado-financeiro-janeiro-de-2022/>>. Citado 2 vezes nas páginas 3 e 10.

SANTOS, C. Estatística descritiva. *Manual de auto-aprendizagem*, v. 2, 2007. Citado na página 5.

Solutis. *Os impactos da pandemia no mercado financeiro*. 2021. Disponível em: <<https://solutis.com.br/2021/05/26/os-impactos-da-pandemia-no-mercado-financeiro/>>. Citado 2 vezes nas páginas 3 e 10.

Anexos

ANEXO A – Código utilizado

```

1 #Janeiro 2019
2 planilha <- read_delim("planilha_201901.csv",
3   delim = ";", escape_double = FALSE, trim_ws = TRUE)
4
5 head(planilha)
6
7 attach(planilha)
8
9 #Substituindo os valores "<=15" da coluna numero_de_operacoes para 15
10 planilha$numero_de_operacoes <- trimws(planilha$numero_de_operacoes, "
   ↳ both")
11 planilha$numero_de_operacoes[planilha$numero_de_operacoes == "<=15"] <-
   ↳ 15
12
13 #convertendo para numerico a coluna numero_de_operacoes
14 planilha$numero_de_operacoes <- as.numeric(planilha$numero_de_operacoes)
15
16 # Cria um grafico de barras interativo com as medias de "numero_de_
   ↳ operacoes" para cada valor unico da coluna "uf"
17 medias_uf <- aggregate(numero_de_operacoes ~ uf, planilha, mean, na.rm =
   ↳ TRUE)
18
19 plot_ly(medias_uf, x = ~uf, y = ~numero_de_operacoes, type = "bar") %>%
20   layout(title = "Media de numero de operacoes por UF", margin = list(b
   ↳ = 150))
21
22 # Cria um grafico de barras interativo com as medias de "numero_de_
   ↳ operacoes" para cada valor unico da coluna "tcb"
23 medias_tcb <- aggregate(numero_de_operacoes ~ tcb, planilha, mean, na.rm
   ↳ = TRUE)
24
25 plot_ly(medias_tcb, x = ~tcb, y = ~numero_de_operacoes, type = "bar")
   ↳ %>%
26   layout(title = "Media de numero de operacoes por TCB", xaxis = list(
   ↳ title = "TCB"), yaxis = list(title = "Media de numero de
   ↳ operacoes"), margin = list(b = 150))
27
28 # Cria um grafico de barras interativo com as medias de "numero_de_
   ↳ operacoes" para cada valor unico da coluna "ocupacao"
29 medias_ocupacao <- aggregate(numero_de_operacoes ~ ocupacao, planilha,
   ↳ mean, na.rm = TRUE)
30

```

```

31 plot_ly(medias_ocupacao, x = ~ocupacao, y = ~numero_de_operacoes, type =
    ↪ "bar") %>%
32   layout(title = "Media_de_numero_de_operacoes_por_ocupacao", margin =
    ↪ list(b = 150))
33
34 # Cria um grafico de barras interativo com as medias de "numero_de_
    ↪ operacoes" para cada valor unico da coluna "cnae_secao"
35 medias_cnae_secao <- aggregate(numero_de_operacoes ~ cnae_secao,
    ↪ planilha, mean, na.rm = TRUE)
36
37 plot_ly(medias_cnae_secao, x = ~cnae_secao, y = ~numero_de_operacoes,
    ↪ type = "bar") %>%
38   layout(title = "Media_de_numero_de_operacoes_por_CNAE_secao", xaxis =
    ↪ list(title = "CNAE_secao"), yaxis = list(title = "Media_de_
    ↪ numero_de_operacoes"), margin = list(b = 150))
39
40 # Cria um grafico de barras interativo com as medias de "numero_de_
    ↪ operacoes" para cada valor unico da coluna "cnae_subclasse"
41 medias_cnae_subclasse <- aggregate(numero_de_operacoes ~ cnae_subclasse,
    ↪ planilha, mean, na.rm = TRUE)
42
43 plot_ly(medias_cnae_subclasse, x = ~cnae_subclasse, y = ~numero_de_
    ↪ operacoes, type = "bar") %>%
44   layout(title = "Media_de_numero_de_operacoes_por_CNAE_subclasse",
    ↪ xaxis = list(title = "CNAE_subclasse"), yaxis = list(title = "
    ↪ Media_de_numero_de_operacoes"), margin = list(b = 150))
45
46 # Cria um grafico de barras interativo com as medias de "numero_de_
    ↪ operacoes" para cada valor unico da coluna "porte"
47 medias_porte <- aggregate(numero_de_operacoes ~ porte, planilha, mean,
    ↪ na.rm = TRUE)
48
49 plot_ly(medias_porte, x = ~porte, y = ~numero_de_operacoes, type = "bar"
    ↪ ) %>%
50   layout(title = "Media_de_numero_de_operacoes_por_porte", xaxis = list(
    ↪ title = "Porte"), yaxis = list(title = "Media_de_numero_de_
    ↪ operacoes"), margin = list(b = 150))
51
52 # Cria um grafico de barras interativo com as medias de "numero_de_
    ↪ operacoes" para cada valor unico da coluna "modalidade"
53 medias_modalidade <- aggregate(numero_de_operacoes ~ modalidade,
    ↪ planilha, mean, na.rm = TRUE)
54
55 plot_ly(medias_modalidade, x = ~modalidade, y = ~numero_de_operacoes,
    ↪ type = "bar") %>%
56   layout(title = "Media_de_numero_de_operacoes_por_modalidade", xaxis =
    ↪ list(title = "Modalidade"), yaxis = list(title = "Media_de_

```

```

    ↪ numero_de_operacoes"), margin = list(b = 150))
57
58 # Cria um grafico de barras interativo com as medias de "numero_de_
    ↪ operacoes" para cada valor unico da coluna "indexador"
59 medias_indexador <- aggregate(numero_de_operacoes ~ indexador, planilha,
    ↪ mean, na.rm = TRUE)
60
61 plot_ly(medias_indexador, x = ~indexador, y = ~numero_de_operacoes, type
    ↪ = "bar") %>%
62   layout(title = "Media_de_numero_de_operacoes_por_indexador", xaxis =
    ↪ list(title = "Indexador"), yaxis = list(title = "Media_de_numero
    ↪ de_operacoes"), margin = list(b = 150))
63
64 # Template para a leitura dos arquivos csv para o banco geral
65 Ano1 <- c(paste0(2019,0,1:9),paste0(2019,10:12))
66
67 Ano2 <- c(paste0(2020,0,1:9),paste0(2020,10:12))
68
69 Ano3 <- c(paste0(2021,0,1:9),paste0(2021,10:12))
70
71 Ano4 <- c(paste0(2022,0,1:9),paste0(2022,10:12))
72
73 BancoGeral <- c(Ano1, Ano2, Ano3, Ano4)
74
75 # Criacao de matrizes vazias com dimensao length(BancoGeral) x 27(numero
    ↪ de estados)
76 matriz_medias_uf_operacoes <- matrix(NA, nrow = length(BancoGeral), ncol
    ↪ = 27,
77                                     dimnames = list(BancoGeral, NULL))
78
79 # Loop for com iteracoes por uf
80 for (j in 1:length(BancoGeral)) {
81   End <- paste0("planilha_", BancoGeral[j], ".csv")
82   planilha <- read_delim(End, delim = ";", escape_double = FALSE, trim_
    ↪ ws = TRUE, show_col_types = FALSE)
83
84   # Substituindo os valores "<=15" da coluna numero_de_operacoes para 15
85   planilha$numero_de_operacoes <- trimws(planilha$numero_de_operacoes, "
    ↪ both")
86   planilha$numero_de_operacoes[planilha$numero_de_operacoes == "<=15"]
    ↪ <- 15
87
88   # Convertendo para numerico a coluna numero_de_operacoes
89   planilha$numero_de_operacoes <- as.numeric(planilha$numero_de_
    ↪ operacoes)
90
91   # Calculando a media da coluna numero_de_operacoes por uf

```

```

92  media_uf_operacoes <- tapply(planilha$numero_de_operacoes, planilha$uf
    ↪ , mean)
93
94  # Armazenando as medias na matriz de medias de operacoes por uf
95  matriz_medias_uf_operacoes[j, ] <- sapply(unique(planilha$uf),
    ↪ function(estado) {
96    replace(media_uf_operacoes[estado], is.na(media_uf_operacoes[estado]
    ↪ ]), 0)
97  })
98
99  }
100
101  # Atribuindo os nomes dos estados as colunas da matriz
102  colnames(matriz_medias_uf_operacoes) <- unique(planilha$uf)
103
104  # Criar objeto ts
105  matriz_medias_uf_operacoes_ts <- ts(matriz_medias_uf_operacoes, start =
    ↪ c(2019, 1), frequency = 12)
106
107  # Definir r tulos das linhas
108  dimnames(matriz_medias_uf_operacoes_ts)[[1]] <- BancoGeral
109
110  # Autoplot
111  autoplot(matriz_medias_uf_operacoes_ts)
112
113  # Estatistica descritiva basica
114  stat.desc(matriz_medias_uf_operacoes_ts)
115
116  # Convertendo para data frame
117  data_frame_uf_operacoes <- as.data.frame(matriz_medias_uf_operacoes)
118
119  # Adicionar coluna de periodo ao data frame
120  data_frame_uf_operacoes <- data_frame_uf_operacoes %>%
121    mutate(Periodo = rep(c("Periodo1", "Periodo2", "Periodo3"), each = 16)
    ↪ )
122
123  # Realizar o teste de Kruskal-Wallis por periodos
124  kruskal_test_uf_operacoes <- kruskal.test(data_frame_uf_operacoes$RS ~
    ↪ data_frame_uf_operacoes$Periodo)
125
126  # Imprimir os resultados
127  print(kruskal_test_uf_operacoes)
128
129  # Teste de dunn
130  dunn_test_uf_operacoes <- dunn.test(data_frame_uf_operacoes$RS, data_
    ↪ frame_uf_operacoes$Periodo, method = "holm")
131

```

```

132 # Estatística descritiva - média e desvio padrão
133 aggregate(data_frame_uf_operacoes$RS ~ data_frame_uf_operacoes$Periodo,
134           ↪ data = data_frame_uf_operacoes, mean)
135
136 aggregate(data_frame_uf_operacoes$RS ~ data_frame_uf_operacoes$Periodo,
137           ↪ data = data_frame_uf_operacoes, sd)
138
139 # Criação da Matriz de médias por Cartão de Crédito PF
140 matriz_medias_cartao_de_credito_pf <- matrix(NA, nrow = length(
141           ↪ BancoGeral), ncol = 10,
142           dimnames = list(BancoGeral,
143           ↪ NULL))
144
145 # Adicionar rótulos às colunas
146 colnames(matriz_medias_cartao_de_credito_pf) <- c("A_vencer_90", "A_
147           ↪ vencer_91_ate_360", "A_vencer_361_ate_1080",
148           ↪ "A_vencer_1081_ate_1800", "A_Vencer_1801_ate_5400", "A_
149           ↪ vencer_acima_5400",
150           ↪ "Vencido_acima_15", "Carteira_Ativa", "Inadimplida_
151           ↪ arrastada", "Ativo_problematICO")
152
153 # Loop for com interações por modalidade Cartão de Crédito - PF #
154           ↪ BancoGeral[j]
155 for (j in 1:length(BancoGeral)) {
156   End <- paste0("planilha_", BancoGeral[j], ".csv")
157   planilha <- read_delim(End, delim = ";", escape_double = FALSE, trim_
158           ↪ ws = TRUE, show_col_types = FALSE)
159
160   # Filtrar as linhas com modalidade igual a "PF - Cartão de Crédito"
161   Names <- unique(planilha$modalidade)
162   planilha_filtrada <- filter(planilha, planilha$modalidade == Names
163           ↪ [1])
164
165   # Convertendo , em .
166   planilha_filtrada <- planilha_filtrada %>%
167     mutate_at(vars(14:23), ~ gsub(",", ".", .))
168
169   # Convertendo para numérico as colunas a vencer/vencidos e outras
170           ↪ relevantes
171   filtro <- apply(planilha_filtrada[,14:23], 2, as.numeric)
172
173   # Teste cálculo mean
174   Mean_filtro <- apply(filtro, 2, mean)
175
176   # Armazenar na matriz_medias_cartao_de_credito as médias do filtro
177   matriz_medias_cartao_de_credito_pf[j,] <- Mean_filtro
178 }

```

```
168
169 # Criar objeto ts
170 matriz_medias_cartao_de_credito_pf_ts <- ts(matriz_medias_cartao_de_
    ↪ credito_pf, start = c(2019, 1), frequency = 12)
171
172 # Definir rotulos das linhas
173 dimnames(matriz_medias_cartao_de_credito_pf_ts)[[1]] <- BancoGeral
174
175 # Autoplot
176 autoplot(matriz_medias_cartao_de_credito_pf_ts)
177
178 # Estatística descritiva básica
179 stat.desc(matriz_medias_cartao_de_credito_pf_ts)
180
181 # Convertendo para data frame
182 data_frame_cartao_de_credito_pf <- as.data.frame(matriz_medias_cartao_de
    ↪ _credito_pf)
183
184 # Adicionar coluna de periodo ao data frame
185 data_frame_cartao_de_credito_pf <- data_frame_cartao_de_credito_pf %>%
186   mutate(Periodo = rep(c("Periodo1", "Periodo2", "Periodo3"), each = 16)
    ↪ )
187
188 # Realizar o teste de Kruskal-Wallis por periodos
189 kruskal_test_cartao_de_credito_pf <- kruskal.test(data_frame_cartao_de_
    ↪ credito_pf$'Carteira Ativa' ~ data_frame_uf_operacoes$Periodo)
190
191 # Imprimir os resultados
192 print(kruskal_test_cartao_de_credito_pf)
193
194 # Teste de dunn
195 dunn_test_cartao_de_credito_pf <- dunn.test(data_frame_cartao_de_credito
    ↪ _pf$'Carteira Ativa', data_frame_cartao_de_credito_pf$Periodo,
    ↪ method = "holm")
196
197 # Estatística descritiva - media e desvio padrao
198 aggregate(data_frame_cartao_de_credito_pf$'Carteira Ativa' ~ data_frame_
    ↪ cartao_de_credito_pf$Periodo, data = data_frame_cartao_de_credito_
    ↪ pf, mean)
199
200 aggregate(data_frame_cartao_de_credito_pf$'Carteira Ativa' ~ data_frame_
    ↪ cartao_de_credito_pf$Periodo, data = data_frame_cartao_de_credito_
    ↪ pf, sd)
```