



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA

O impacto do Efeito da Idade Relativa no desempenho de jogadores de futebol: Um estudo com base nos dados da Série A do Campeonato Brasileiro de 2023

Autor: Giordano Maciel da Silva
Orientadora: Prof. Dra. Luciana Neves Nunes

Porto Alegre
2024

GIORDANO MACIEL DA SILVA

O impacto do Efeito da Idade Relativa no desempenho de jogadores de futebol: Um estudo com base nos dados da Série A do Campeonato Brasileiro de 2023

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do título de Bacharel em Estatística, em curso de graduação oferecido pela Universidade Federal do Rio Grande do Sul.

Orientadora: Professora Dra. Luciana Neves Nunes, UFRGS

Banca Examinadora:

Professora Dra. Lisiane Priscila Roldão Selau, UFRGS

Professor Dr. Marcio Valk, UFRGS

Porto Alegre
2024

Dedico este trabalho a todos aqueles que vêm o futebol para além do esporte, mas também, como um poderoso meio de união entre povos, uma ferramenta de transformação de vidas e alento para os momentos difíceis.

Também, para todos aqueles que fazem do futebol um objeto de estudo, e que o vêm como algo que pode estar diretamente ligado à ciência e à pesquisa.

Por fim, dedico a minha maior incentivadora, e, agora, meu anjo, Elisângela Caetano Maciel (In Memoriam).

AGRADECIMENTOS

Ao meu pai, Carlos, que me introduziu ao mundo do futebol e cultivou em mim a paixão por esse esporte desde a infância. À minha mãe, Vera, por sempre acreditar na minha capacidade de aprender, incentivando-me nos estudos e a dedicar-me ao que amo.

À minha irmã, Carla, minha maior inspiração na vida acadêmica, que abriu as portas das universidades para nossa família e pavimentou o caminho que também percorri. Além disso, por ser minha companheira de residência durante a graduação e por auxiliar na construção deste trabalho.

A todos os familiares e amigos de infância e adolescência por permanecerem presentes, mesmo à distância. O apoio e incentivo de cada um foram fundamentais, seja nas conversas pessoais ou virtuais.

À UFRGS e ao Instituto de Matemática e Estatística, por proporcionarem uma educação pública de qualidade.

À minha orientadora, Luciana, por acreditar neste trabalho desde o início, compreendendo a importância que esse tema tem para mim.

Aos colegas de graduação, especialmente aos meus cinco companheiros desde a pandemia até o retorno presencial: Antônio, Kevin, João Lucas, Leonardo e Leticia. Por trazerem leveza até para os momentos mais desafiadores desta etapa da vida.

A todos vocês, o meu mais sincero agradecimento por representarem tanto neste momento. O estatístico que sou hoje tem uma parte significativa de vocês nessa jornada. O meu eu criança, que registrava números e criava tabelas dos esportes e programas favoritos, já utilizava a estatística sem compreender, e hoje tenho a honra de entender o papel crucial dessa ciência em nossa sociedade.

RESUMO

Este trabalho investigou o impacto do Efeito da Idade Relativa (EIR) no desempenho de jogadores de futebol na Série A do Campeonato Brasileiro de 2023. Para as análises estatísticas foram utilizados o teste qui-quadrado de aderência e a análise de variância (ANOVA). Os resultados confirmam a presença significativa do EIR, principalmente entre jogadores nascidos nas décadas de 1990 e 2000, mostrando que o trimestre de nascimento pode ser um fator determinante para a ascensão dos jogadores das categorias de base para o nível profissional. A análise temporal revelou variações, sendo mais evidente nas décadas mencionadas em comparação com a década de 1980. Observou-se uma influência diferenciada nas posições de campo, com o EIR mais acentuado entre Defensores, Meias e Atacantes, enquanto não se manifestou entre goleiros. No aspecto do desempenho, jogadores nascidos no 4º Trimestre mostraram médias superiores em várias métricas, especialmente entre os Defensores. Essas descobertas sugerem implicações práticas relevantes para treinadores, gestores e olheiros no futebol, destacando a necessidade de considerar o Efeito da Idade Relativa ao avaliar e desenvolver jovens talentos. Este estudo pode contribuir para a literatura ao fornecer insights sobre o Efeito da Idade Relativa no contexto do futebol brasileiro, abrindo caminho para investigações futuras que aprofundem a compreensão desse fenômeno no ambiente esportivo. Conclui-se que, para maximizar o desenvolvimento de jovens talentos, estratégias devem ser implementadas para mitigar o impacto do EIR nos processos de escolha das categorias de base.

Palavras-chave: Efeito da Idade Relativa, Futebol Brasileiro, Categorias de base, Qui-quadrado de aderência

ABSTRACT

This work investigated the impact of the Relative Age Effect (RAE) on the performance of football players in Series A of the 2023 Brazilian Championship. The chi-square test of adherence and analysis of variance (ANOVA) were used for statistical analyses. The results confirm the significant presence of RAE, especially among players born in the 1990s and 2000s, showing that the 4th quarter of birth can be a determining factor in the rise of players from the youth categories to the professional level. The temporal analysis revealed variations, being more evident in the mentioned decades compared to the 1980s. A different influence was observed in the field positions, with the RAE being more pronounced among Defenders, Midfielders and Forwards, while it was not evident among goalkeepers. In terms of performance, players born in the 4th Quarter showed superior averages in several metrics, especially among Defenders. These findings suggest relevant practical implications for football coaches, managers and scouts, highlighting the need to consider the Relative Age Effect when evaluating and developing young talent. This study can contribute to the literature by providing insights into the Relative Age Effect in the context of Brazilian football, paving the way for future investigations that deepen our understanding of this phenomenon in the sporting environment. It is concluded that, to maximize the development of young talent, strategies must be implemented to mitigate the impact of the RAE on the selection processes for youth teams.

Key-words: Relative Age Effect, Brazilian Soccer, Youth team, Chi-square test of adherence

Lista de Figuras

1. Cronograma da Metodologia	15
2. Distribuição de todos os jogadores do campeonato por trimestre de nascimento	23
3. Distribuição de jogadores com EIR por trimestre de nascimento	27
4. <i>Boxplots</i> para a proporção média de participação no campeonato	29
5. <i>Boxplots</i> para a proporção média de titularidade no campeonato	29
6. <i>Boxplots</i> para a média de minutos jogados no campeonato	29
7. <i>Boxplots</i> para a média de minutos jogados por jogo disputado no campeonato	30
8. <i>Boxplots</i> para a média da Nota <i>SofaScore</i> no campeonato	30

Lista de Tabelas

1. Frequência de Jogadores do campeonato por posição	21
2. Frequência de Jogadores do campeonato por nacionalidade	21
3. Frequência de Jogadores do campeonato por década de nascimento	22
4. População de referência	22
5. Teste qui-quadrado de aderência em todos os jogadores do campeonato	23
6. Teste qui-quadrado de aderência em todos os jogadores do campeonato - década de nascimento	24
7. Teste qui-quadrado de aderência em todos os jogadores do campeonato - posição	25
8. Teste qui-quadrado de aderência para os defensores - década de nascimento	25
9. Teste qui-quadrado de aderência para os meias - década de nascimento	26
10. Teste qui-quadrado de aderência para os atacantes - década de nascimento	26
11. Teste qui-quadrado de aderência para os jogadores com EIR	27
12. Distribuição dos jogadores que estiveram em todo o campeonato	28
13. Quadro resumo das variáveis quantitativas por Trimestre	28
14. Quadro resumo das variáveis quantitativas por Trimestre - Defensores	30
15. Quadro resumo das variáveis quantitativas por Trimestre - Meias	31
16. Quadro resumo das variáveis quantitativas por Trimestre - Atacantes	31

1 INTRODUÇÃO	9
2 REFERENCIAL TEÓRICO	11
2.1 Efeito da Idade Relativa	11
2.2 Análise de desempenho em jogadores de futebol	13
3 METODOLOGIA	15
3.1 Coleta dos dados e medidas descritivas	15
3.2 Teste qui-quadrado de aderência e resíduos padronizados	16
3.2.1 Estatística de teste	17
3.2.2 Resíduos padronizados	17
3.3 Análise de variância (ANOVA) e Complemento de Tukey	19
3.3.1 Estatística de teste ANOVA	19
3.3.2 Complemento de Tukey para Comparação Múltipla	19
4 RESULTADOS E DISCUSSÃO	21
4.1 Análises Iniciais	21
4.2 Grupos de jogadores com a presença do EIR	26
4.3 Grupos de jogadores para a análise quantitativa	27
5 CONCLUSÕES	32
5.1 Efeito da Idade Relativa: Evidências e Variações	32
5.2 Tendências de Desempenho	33
5.3 Limitações e Perspectivas para pesquisas futuras	33
5.4 Contribuições e Relevância	34
REFERÊNCIAS	

1 INTRODUÇÃO

A prática do esporte, notadamente o futebol, é um fenômeno de alcance global, que envolve milhões de pessoas, desde as categorias de base até o nível profissional. A busca por jovens talentos promissores é uma constante muito importante nas academias e clubes (Cruz, 2012), e, nos últimos anos, a análise dos fatores que influenciam o sucesso dos jogadores no mundo do futebol tem ganhado destaque no âmbito acadêmico. Um desses fatores, denominado "idade relativa", diz respeito à influência da data de nascimento de um jogador em relação aos seus pares dentro de uma mesma categoria de base. A diferença de idade, mesmo que seja apenas de meses, durante a adolescência, pode resultar em vantagens significativas em termos de condições físicas, cognitivas e psicológicas para os jogadores nascidos nos primeiros meses do ano. Essas vantagens podem ter um impacto positivo em seu desempenho esportivo (Musch e Grondin, 2001; Massa et al, 2020).

A teoria do efeito da idade relativa ganhou destaque na literatura científica e esportiva, com estudos anteriores demonstrando sua existência em diferentes esportes, como hóquei no gelo (Barnsley et al., 1985), basquete (Delorme e Raspaud, 2009) e, inclusive, no futebol (Williams, 2010). O trabalho presente tem como tema central "O Efeito da Idade Relativa no Futebol". Ele aborda a hipótese de que jogadores de futebol nascidos em determinados trimestres do ano podem apresentar vantagens ou desvantagens em relação aos seus colegas devido ao calendário de competições adotado nas categorias de base.

Enquanto os atletas estão atuando nas categorias de base, eles têm a oportunidade de desenvolver os principais atributos necessários para se tornarem jogadores profissionais no futebol. Além da força física, eles aprimoram sua capacidade intelectual, cognitiva e outras habilidades essenciais, como agilidade, técnica refinada, resistência e trabalho em equipe (Massa, 2017). No entanto, surge uma questão relevante: a idade relativa e seu possível impacto significativo na probabilidade de um jogador alcançar o nível profissional. Em outras palavras, será que jogadores nascidos em determinados trimestres do ano possuem uma vantagem em relação àqueles nascidos em outros trimestres do mesmo ano quando se trata de atingir o futebol profissional?

Levando em conta essas informações, o trabalho tem como objetivo principal analisar o efeito da idade relativa no futebol a partir de dados do Campeonato Brasileiro de 2023, investigando se jogadores nascidos em diferentes trimestres do ano apresentam diferenças na probabilidade de alcançar o nível profissional. Além disso, busca-se determinar se existe uma diferença significativa no desempenho dos jogadores nascidos em diferentes trimestres quando eles já estão no futebol profissional, avaliando possíveis diferenças no comportamento das métricas pelas posições e décadas de nascimento dos jogadores

Para desenvolver o tema, o estudo é iniciado com esta introdução que aborda a importância das categorias de base e a influência do efeito da idade relativa na formação de jogadores. No segundo capítulo, teremos o referencial teórico que fundamentará a pesquisa, abordando os conceitos de Efeito da Idade Relativa; Análise de desempenho em jogadores de futebol; e Teste Qui-Quadrado de Aderência.

Ainda, no terceiro capítulo, será abordada a base metodológica de desenvolvimento do trabalho, constando quais serão as técnicas utilizadas para análises e como o banco de dados do estudo foi construído. Após esta seção, adentraremos ao capítulo de Resultados e Discussões, destacando as principais descobertas deste estudo, seguido pelas conclusões que oferecerão uma síntese acerca do tema explorado nesta pesquisa.

Por meio deste trabalho, espera-se obter informações que possam contribuir para aprimorar os processos de seleção e desenvolvimento de talentos no futebol, promovendo maior equidade e justiça para todos os jogadores das categorias de base, independentemente do trimestre de nascimento.

2 REFERENCIAL TEÓRICO

2.1 Efeito da Idade Relativa

O "Efeito da Idade Relativa" (EIR) é um fenômeno amplamente estudado no campo da psicologia do desenvolvimento, que descreve a tendência de que indivíduos nascidos no início de um ano calendário tenham vantagens comparativas em relação aos nascidos nos meses posteriores do mesmo ano. Esse fenômeno, que está intimamente relacionado à forma como as datas de corte são estabelecidas em diferentes contextos, tem sido objeto de investigação em diversas áreas, incluindo educação, esportes e desenvolvimento humano.

O EIR, conforme descrito por Marques, Pinheiro e Coswig (2018), evidencia que atletas nascidos nos primeiros meses do ano contam com vantagens decorrentes de sua maior idade cronológica em relação aos nascidos no final do ano. Essa disparidade, reconhecida como um fenômeno notável no âmbito esportivo, tem raízes na década de 1960, quando estudos realizados por Dickinson e Larson (1963) e Guily (1965) inicialmente revelaram a influência dos meses de nascimento no desempenho escolar de crianças e adolescentes.

A descoberta do EIR na área educacional desencadeou um interesse crescente na sua aplicação em contextos esportivos. Barnsley e Thompson (1985) ampliaram essa investigação ao explorar a relação entre o mês de nascimento e a probabilidade de atletas se tornarem profissionais. Em um estudo subsequente, Barnsley e Thompson (1988) aprofundaram a análise, focando especificamente em atletas de hóquei da National Hockey League. Suas pesquisas revelaram que atletas nascidos nos dois primeiros trimestres do ano apresentavam maior probabilidade de ascender ao nível de atletas de alto desempenho. Esses achados reforçaram a noção de que a vantagem relacionada ao mês de nascimento não apenas afeta o desempenho acadêmico, mas também desempenha um papel crucial no desenvolvimento e sucesso atlético, tornando o EIR um fenômeno multidimensional com implicações significativas em diversas áreas da vida.

No contexto esportivo, o EIR desempenha um papel significativo, e o futebol não é uma exceção. Desde tenra idade, as crianças envolvidas em programas de futebol, baseados em categorias etárias, enfrentam o desafio do EIR. Jogadores mais velhos, especialmente aqueles que nasceram nos primeiros meses do ano, frequentemente apresentam vantagens

físicas e psicológicas sobre seus colegas mais jovens. Isso pode resultar em maior tempo de jogo, mais oportunidades de treinamento de alto nível e, conseqüentemente, um desenvolvimento mais acelerado em relação aos jogadores mais jovens.

Em um estudo conduzido por Hirose (2009), observou-se que mais da metade (58,8%) dos jogadores de futebol do Japão nasceram no primeiro trimestre do ano, enquanto apenas 24% da população japonesa nasceu nesse período do ano. Essa discrepância revela a influência do EIR na composição demográfica dos jogadores de futebol, destacando a prevalência de nascimentos nos primeiros meses do ano no cenário do futebol japonês.

O EIR também tem sido investigado quanto aos seus efeitos hormonais, parâmetros antropométricos e desempenho físico em jovens jogadores de futebol, como apontado por estudos como o de Moreira et al. (2013). Essas análises ressaltam a complexidade do processo de treinamento e desenvolvimento, com o EIR influenciando não apenas fatores psicológicos, mas também variáveis fisiológicas essenciais para o progresso esportivo.

Além disso, a identificação de talentos no futebol, um elemento vital para clubes de elite, é afetada pelo EIR, conforme discutido por Williams (2010). Em uma análise das edições da Copa do Mundo de Futebol Sub-17 da FIFA entre 1997 e 2007, Williams observou uma distribuição significativamente diferente dos meses de nascimento, com uma predominância de jogadores nascidos nos primeiros meses do ano. Essa tendência impacta as decisões de recrutamento, evidenciando a importância de estratégias mais equitativas para identificar talentos, independentemente do mês de nascimento.

Recentemente, Neto et al. (2020) exploraram a manifestação do Efeito da Idade Relativa (EIR) no desempenho dos jogadores que participaram do Campeonato Brasileiro de 2019. Os resultados deste estudo indicaram a ausência de diferenças significativas no desempenho dos atletas em relação aos trimestres de nascimento. No entanto, observou-se uma relação entre tempo de jogo e o trimestre de nascimento, indicando que os jogadores nascidos nos últimos meses do ano teriam mais minutos disputados do que aqueles nascidos nos primeiros meses do ano. Diante desse contexto, a presente pesquisa, focada no Campeonato Brasileiro de 2023, busca aprofundar a análise, incorporando mais variáveis, além do tempo de jogo, para verificar se essa tendência persiste no cenário do futebol brasileiro.

2.2 Análise de desempenho em jogadores de futebol

A prática de *scouts* ou análises estatísticas no futebol tem raízes no século XIX, porém, é a partir de publicações como o livro "Os Números do Jogo: Por que Tudo o que Você Sabe sobre o Futebol Está Errado" (Anderson e Sally), de 2013, que essa abordagem atingiu um marco significativo ao introduzir a discussão sobre análises numéricas no esporte.

A aplicação de *scouts* não apenas eleva a análise do jogo para os espectadores, mas também traz benefícios expressivos para as equipes. Ao coletar, processar e analisar informações, as comissões técnicas conseguem avaliar o desempenho coletivo e individual, identificar variáveis que influenciam o rendimento, corrigir falhas e aprimorar os pontos fortes dos jogadores. Essas análises, que abrangem áreas como passes, finalizações, desarmes e roubadas de bola, desempenham um papel crucial na formulação de estratégias, táticas e até mesmo nas decisões de contratação.

A atual revolução no cenário do futebol destaca a crescente importância da análise estatística para compreender o desempenho dos jogadores. Anteriormente limitado ao que os torcedores viam na televisão, esse cenário passou por uma mudança considerável, com aproximadamente 1.500 eventos distintos a cada rodada, exigindo uma abordagem mais quantitativa.

Nesse contexto, a *SofaScore*¹, empresa dedicada ao agrupamento e publicações de dados esportivos, emerge como uma das respostas a essa necessidade, criando uma pontuação de desempenho para os jogadores baseados nas ações dentro de campo, empregando operadores humanos na coleta manual de dados sobre eventos como posses, passes, duelos, desarmes e chutes. A abordagem busca traduzir essas estatísticas em informações úteis e claras que resumem o desempenho do jogador.

A escala de classificação de scores, variando de 6 a 10, proporciona uma interpretação rápida da performance. Pontuações abaixo de 6 são raras, indicando erros significativos, enquanto uma pontuação perfeita de 10 é reservada para esforços extraordinários com erros mínimos. O algoritmo *SofaScore* busca uma avaliação equitativa, considerando a

¹ <https://www.sofascore.com/pt/>

complexidade do futebol, permitindo que jogadores com habilidades diversas recebam pontuações mais altas, promovendo uma visão imparcial do desempenho.

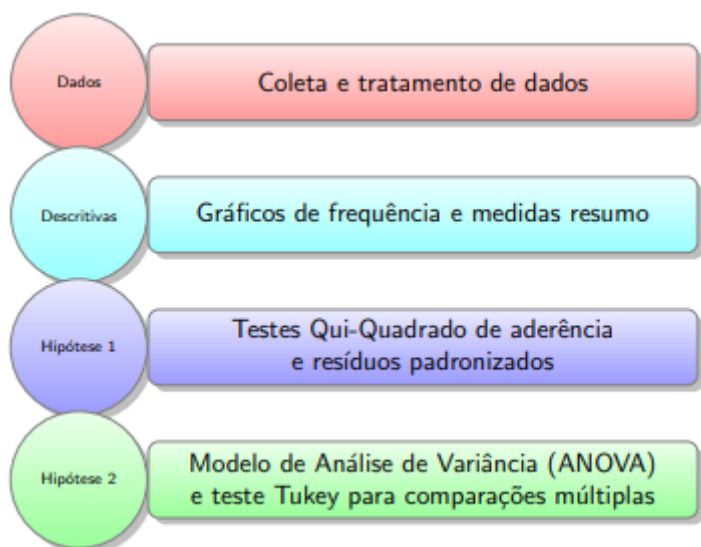
O futebol está passando por uma transformação em direção a se tornar um produto orientado por dados, destacando a importância das técnicas de avaliação do desempenho. Essa evolução não apenas representa uma mudança na análise, mas também uma revolução na compreensão do jogo, enfatizando a importância de métricas objetivas na avaliação individual e coletiva (Shamah, 2021).

3 METODOLOGIA

Esta pesquisa adota uma abordagem quantitativa, utilizando dados secundários provenientes de fontes confiáveis e acessíveis relacionadas ao futebol profissional no Brasil. O tratamento do banco de dados e a construção dos códigos do testes e análises foi feito por meio do software livre R, utilizando o RStudio como ferramenta de visualização. A fim de ajudar nas análises e visualizações dos resultados foram utilizados alguns pacotes, como: kable, dplyr, readxl, knitr, kableExtra, ggplot2, etc. As demais funções utilizadas fazem parte das funções previamente instaladas na linguagem R.

Para a realização do trabalho foram seguidas quatro etapas, conforme Figura 1.

Figura 1. Cronograma da Metodologia



3.1 Coleta dos dados e Medidas descritivas

Para a construção do banco de dados, variáveis como idade, time, nacionalidade e trimestre de nascimento foram extraídos do site *Football Reference*², enquanto informações como Nota de Desempenho, Minutos Jogados e Jogos Disputados foram obtidas do site *SofaScore*. A coleta foi realizada em 07/12/2023, um dia após a conclusão da última rodada do campeonato. O banco de dados coletado inclui informações de 732 jogadores do Campeonato Brasileiro de Futebol Série A - 2023.

² <https://fbref.com/pt/>

As variáveis de interesse abrangem uma variedade de fatores relevantes para o estudo, incluindo idade atual, década, ano e trimestre de nascimento, nacionalidade, posição do jogador no campo de jogo e variáveis associadas ao desempenho no futebol. Na variável posição, adotamos a classificação que o site *SofaScore* utiliza, sendo: Goleiros, Defensores (zagueiros e laterais), Meias (volantes e armadores), Atacantes (pontas e centroavantes).

Considerando que o efeito estudado pode não ser uniforme em todo o mundo, excluimos da análise jogadores que não realizaram suas categorias de base na América do Sul. Além disso, 10 jogadores foram excluídos do banco de dados por não terem nota atribuída no site *SofaScore*, devido a terem participado de apenas uma partida, entrando nos minutos finais do jogo. Ao final, permaneceram 712 jogadores.

Antes da realização da análise estatística, procedeu-se à coleta de informações sobre a população de referência do estudo. Essas informações foram obtidas no sistema *DataSus*³, em 16/12/2023. Foram registrados os números de nascidos vivos na população masculina no período de 1994 a 2006, uma escolha justificada pelo fato de que 1994 é o ano mais antigo disponível no sistema, e 2006 representa o ano de nascimento do jogador mais jovem presente no banco de dados.

A análise estatística empregou diversas técnicas. Para uma exploração inicial dos dados, utilizamos gráficos de frequência. Variáveis contínuas foram resumidas por meio de estatísticas como média com intervalos de confiança de 95% e amplitude interquartílica resumida em *boxplots*. Já as variáveis qualitativas foram descritas em termos de frequências absolutas e relativas (percentual).

3.2 Teste qui-quadrado de aderência e resíduos padronizados

O teste Qui-Quadrado, conhecido como Qui-Quadrado de Pearson, é uma ferramenta estatística amplamente utilizada na análise de dados categóricos e comparação de proporções em estudos de pesquisa (Agresti, 2007). Foi inicialmente desenvolvido por Karl Pearson em 1900 (Pearson, 1900) e desde então tem se estabelecido como um método fundamental em diversas áreas de pesquisa.

³ <http://tabnet.datasus.gov.br/>

3.2.1 Estatística de teste

O modelo do teste Qui-Quadrado de aderência envolve a criação de uma tabela com a variável categórica e a população de referência, denominada de “grupo lei”. O teste Qui-Quadrado calcula a estatística de Qui-Quadrado χ^2 com base nas frequências observadas e esperadas nas células da tabela. A estatística χ^2 é usada para avaliar se existe uma não-aderência significativa entre as proporções das categorias da variável e da população de referência. O valor crítico do Qui-Quadrado é comparado ao valor calculado para determinar se a não-aderência entre as populações é estatisticamente significativa.

Para calcular a estatística χ^2 , utilizamos a seguinte fórmula:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Em que:

- O representa as frequências observadas.
- E representa as frequências esperadas.
- A soma é feita para todas as células da tabela de contingência.

A interpretação do teste Qui-Quadrado envolve a comparação do valor calculado de χ^2 com um valor crítico disponível na tabela da distribuição Qui-Quadrado, geralmente com base em um nível de significância pré-definido (por exemplo, 0,05). Se o valor calculado de χ^2 for maior do que o valor crítico, rejeita-se a hipótese nula, indicando que existe uma associação significativa entre as variáveis. Por outro lado, se o valor calculado for menor que o valor crítico, não há evidências suficientes para rejeitar a hipótese nula, sugerindo que as proporções nas variáveis não são significativamente diferentes.

3.2.2 Resíduos Padronizados

Além da estatística χ^2 no teste Qui-Quadrado, outra medida relevante na análise de dados categóricos é a utilização de resíduos padronizados. Agresti (2007) destaca a importância desses resíduos como uma ferramenta adicional para compreender a significância

estatística das células individuais na tabela de contingência, que pode ser chamada de “associação local”.

Os resíduos padronizados são calculados para cada célula da tabela e representam a diferença entre a frequência observada (O) e a frequência esperada (E), dividida pelo desvio padrão da célula. Matematicamente, a fórmula para o cálculo do resíduo padronizado (Z) é expressa por:

$$Z = \frac{(O - E)}{\sqrt{E}}$$

Em que:

- O representa as frequências observadas.
- E representa as frequências esperadas.

Ao calcular os resíduos padronizados para cada célula, obtém-se uma medida de quão longe as frequências observadas estão daquelas que seriam esperadas sob a hipótese nula de aderência à população esperada.

Resíduos padronizados significativamente diferentes de zero indicam que a célula específica contribui de maneira substancial para a diferença global observada nas proporções das variáveis, ou seja, há evidência de que exista associação local naquela célula. Essa análise mais detalhada das células ressalta áreas específicas de associação ou discordância, fornecendo *insights* adicionais sobre as relações entre as variáveis. Considerando nível de significância de 5% ($\alpha=0,05$), o valor crítico do teste é 1,96. Sendo um teste bilateral, considera-se que quando o valor da estatística do teste é menor que -1,96, a associação é negativa, isto é, o valor observado é significativamente menor que o esperado, sob a hipótese nula. Já quando o valor é maior que 1,96, tem-se que a associação é positiva, isto é, o valor observado é significativamente maior que o esperado, sob a hipótese nula.

Em suma, os resíduos padronizados no contexto do teste Qui-Quadrado oferecem uma abordagem mais refinada para avaliar a significância estatística em nível celular, aprimorando

a capacidade de identificar associações específicas entre as categorias das variáveis categóricas em análise.

3.3 Análise de Variância (ANOVA) e Complemento de Tukey

A Análise de Variância (ANOVA) é uma técnica estatística utilizada para comparar as médias de três ou mais grupos independentes. Ela é uma extensão do teste t de Student, que compara as médias de apenas dois grupos. A ANOVA testa a hipótese nula de que não há diferença significativa entre as médias dos grupos e a hipótese alternativa de que pelo menos uma média é significativamente diferente das outras.

3.3.1 Estatística de Teste na ANOVA

Na ANOVA, a estatística de teste é calculada comparando a variabilidade entre os grupos com a variabilidade dentro dos grupos. Se a variabilidade entre os grupos for significativamente maior do que a variabilidade dentro dos grupos, isso sugere que as médias dos grupos são diferentes de forma estatisticamente significativa.

A estatística de teste F na ANOVA é calculada dividindo a variabilidade entre os grupos pela variabilidade dentro dos grupos. Matematicamente, a fórmula para a estatística F é expressa por:

$$F = \frac{MS \text{ entre}}{MS \text{ dentro}}$$

Em que:

- MS entre é a média dos quadrados entre os grupos, calculada como a soma dos quadrados entre os grupos dividida pelo número de grupos menos 1.
- MS dentro é a média dos quadrados dentro dos grupos, calculada como a soma dos quadrados dentro dos grupos dividida pelo total de graus de liberdade dentro dos grupos.

3.3.2 Complemento de Tukey para Comparação Múltipla

Após encontrar uma diferença significativa entre os grupos utilizando a ANOVA, é comum realizar testes adicionais para identificar quais grupos específicos diferem entre si.

Uma abordagem popular para essas comparações múltiplas é o complemento de Tukey, também conhecido como teste de Tukey ou procedimento de Tukey.

O complemento de Tukey é um método estatístico utilizado para comparar todas as possíveis combinações de médias de grupos e determinar quais são estatisticamente diferentes. Ele controla o erro do Tipo I (rejeitar erroneamente a hipótese nula) ao realizar múltiplas comparações, ajustando os intervalos de confiança para cada comparação.

O teste de Tukey calcula um intervalo de confiança para a diferença entre cada par de médias e compara esse intervalo com uma quantidade crítica, que depende do número total de grupos e do tamanho da amostra. Se a diferença entre duas médias for maior do que o intervalo de confiança, essas médias são consideradas significativamente diferentes.

Em resumo, a ANOVA é utilizada para determinar se há diferenças significativas entre os grupos, enquanto o complemento de Tukey é empregado para identificar quais grupos diferem entre si após a detecção de uma diferença global.

4 RESULTADOS E DISCUSSÃO

4.1 Análises Iniciais

Nas três primeiras tabelas, são apresentadas informações detalhadas sobre a estrutura do banco de dados, incluindo as posições dos jogadores, suas nacionalidades e as décadas de seus nascimentos. É importante ressaltar que nosso banco de dados é composto por 712 jogadores, após a exclusão de 20 entradas iniciais. Desses 712, podemos perceber que a maioria se concentra nas posições de linha do campo (Defensores, Meias e Atacantes), com apenas 47 sendo goleiros (Tabela 1).

Tabela 1. Frequência de Jogadores do campeonato por posição

Posição	Frequência absoluta	Frequência relativa
Goleiros	47	6.6%
Defensores	230	32.3%
Meias	271	38%
Atacantes	164	23.1%

É relevante destacar que, dentre os jogadores analisados, 406 (57,1%) nasceram na década de 1990, tornando esta a década predominante entre os atletas (Tabela 2). Além disso, é notável que a expressiva maioria, totalizando 599 (84,25%), possui nacionalidade brasileira (Tabela 3). Esses dados proporcionam uma visão abrangente da composição do banco de dados, evidenciando a representatividade significativa de jogadores brasileiros e a concentração temporal na década de 1990.

Tabela 2. Frequência de Jogadores do campeonato por década de nascimento

Década de Nascimento	Frequência absoluta	Frequência relativa
1970	1	0.1%
1980	103	14.5%
1990	406	57%
2000	202	28.4%

Tabela 3. Frequência de Jogadores do campeonato por nacionalidade

Nacionalidade	Frequência absoluta	Frequência relativa
Brasil	599	84.2%
Argentina	37	5.2%
Uruguai	28	4%
Colômbia	14	2%
Paraguai	12	1.6%
Equador	9	1.3%
Chile	7	1.1%
Venezuela	4	0.5%
Bolívia	1	0.1%

Conforme o conceito do efeito da idade relativa, é crucial estabelecer uma população de referência para a devida comparação com a nossa amostra. Nesse sentido, utilizando o sistema *DataSus*, foram coletados os registros de nascidos vivos no período compreendido entre 1994 e 2006. Ressalta-se que, para os propósitos desta pesquisa, foram tabulados exclusivamente os dados referentes à população masculina, uma vez que a análise está centrada no contexto do Campeonato Brasileiro de Futebol Masculino. A partir desses procedimentos, obtivemos a distribuição da população de referência para este estudo, disponível na Tabela 4.

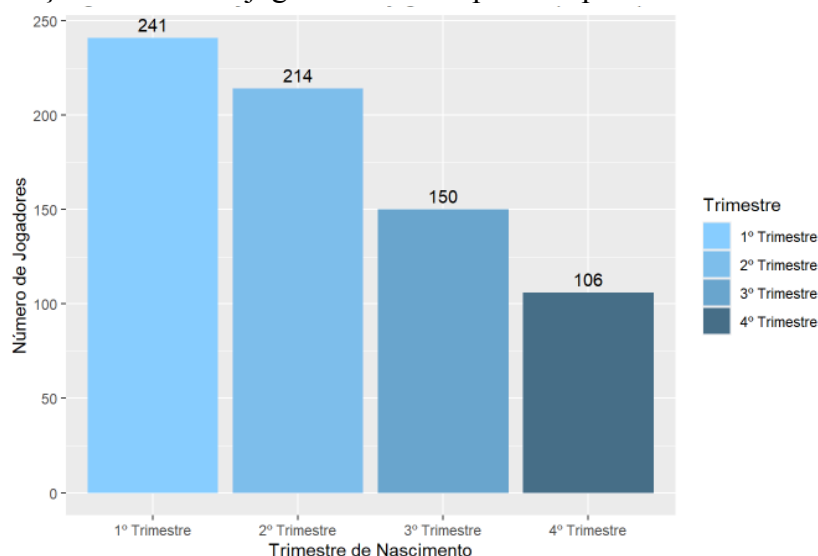
Tabela 4. População de referência com informações coletadas no site do Datasus

	População BR Masc 1994-2006	Distribuição em %
1º Trimestre	4.907.001	25.6%
2º Trimestre	5.059.066	26.4%
3º Trimestre	4.762.577	24.8%
4º Trimestre	4.439.898	23.2%
Total	19.168.542	100%

Fonte: DataSus (<http://tabnet.datasus.gov.br/>).

Com a distribuição da população de referência em mãos, procedemos à análise e comparação desta com os dados de nosso banco. Na Figura 2, é possível visualizar a distribuição dos nascimentos por trimestre, revelando uma tendência linear. Destaca-se que há uma predominância de jogadores nascidos no primeiro trimestre (241; 33,9%), superando as quantidades nos trimestres subsequentes (214; 30%, 150; 21,1%, 106; 15%).

Figura 2. Distribuição de todos os jogadores do campeonato por trimestre de nascimento



De acordo com a Tabela 5, logo abaixo, pode-se ver que há evidência de que a distribuição de nascimentos dos jogadores não tem aderência a distribuição de nascimentos da população em geral ($p < 0,001$). Quanto aos resíduos, é possível se observar que nos dois primeiros trimestres, observa-se um número de jogadores significativamente superior ao esperado (Resíduo $> 1,96$), enquanto nos dois últimos trimestres, o número é significativamente inferior ao esperado (Resíduo $< -1,96$).

Esses resultados indicam uma disparidade estatisticamente significativa entre a distribuição de nascimentos na população de referência e a composição do nosso banco de dados, evidenciando uma possível influência do efeito da idade relativa no perfil dos jogadores analisados.

Tabela 5. Teste qui-quadrado de aderência em todos os jogadores do campeonato

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total	χ^2	P-valor
População BR*	25.6%	26.4%	24.8%	23.2%	100%		
Jogadores do BR 23	241 (33.9%)	214 (30%)	150 (21.1%)	106 (15%)	711 (100%)	46.851	< 0.001
Resíduos	5.001	2.237	-2.286	-5.182			

*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

Dado que o cenário do futebol, incluindo as categorias de base e os processos de seleção, passa por evoluções ao longo do tempo, é razoável considerar a hipótese de que o efeito da idade relativa não seja uniforme em diferentes décadas. Diante desse contexto, optamos por agrupar os jogadores de acordo com a década de seu nascimento (1980, 1990 e 2000).

Cabe ressaltar que um jogador nascido na década de 1970 será excluído dessa análise devido à insuficiência de dados para uma análise robusta dessa década. Além disso, é importante notar que a década de 2000 abrange até o ano de 2006, sendo este o ano de nascimento do jogador mais jovem que disputou o Campeonato Brasileiro de 2023.

Na Tabela 6, apresentamos testes qui-quadrado de aderência para as décadas analisadas. Os resultados indicam que nas décadas de 1990 e 2000 existem indícios da presença do Efeito da Idade Relativa ($p < 0,001$), ao passo que na década de 1980, a evidência sugere a ausência desse efeito ($p = 0,6712$). Essa diferenciação destaca a importância de considerar a variabilidade temporal ao investigar o impacto do EIR nas diferentes épocas do futebol.

Tabela 6. Teste qui-quadrado de aderência em todos os jogadores do campeonato - década de nascimento

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total	χ^2	P-valor
População BR*	25.6%	26.4%	24.8%	23.2%	100%		
Década de 1980	21 (20.4%)	29 (28.1%)	25 (27.2%)	25 (24.3%)	106 (100%)	1.5483	0.6712
Resíduos	-1.233	0.404	0.560	0.282			
Década de 1990	142 (35.1%)	120 (29.6%)	84 (20.7%)	59 (14.6%)	405 (100%)	30.865	<0.001
Resíduos	4.311	1.474	-1.891	-4.073			
Década de 2000	78 (38.6%)	65 (32.2%)	37 (18.3%)	22 (10.9%)	202 (100%)	32.12	<0.001
Resíduos	4.200	1.863	-2.133	-4.117			

*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

Também realizamos análises específicas para as posições dos jogadores no campeonato, uma vez que a literatura sustenta que o Efeito da Idade Relativa (EIR) pode manifestar-se em determinadas posições e não necessariamente em outras. A partir dos resultados apresentados na Tabela 7, podemos inferir a presença do EIR entre os Defensores ($p < 0,001$), Meias ($p < 0,001$) e Atacantes ($p = 0,043$). No entanto, observa-se uma não-significância para os goleiros, uma vez que o p-valor do teste foi de 0,4148.

Esses achados reforçam a ideia de que o EIR não se manifesta uniformemente em todas as posições do campo de jogo. Essa diferenciação ressalta a importância de uma abordagem segmentada ao avaliar o impacto do Efeito da Idade Relativa, reconhecendo suas variações de acordo com a posição ocupada pelos jogadores no contexto do campeonato.

Tabela 7. Teste qui-quadrado de aderência em todos os jogadores do campeonato - posição

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total	χ^2	P-valor
População BR*	25.6%	26.4%	24.8%	23.2%	100%		
Goleiros	15 (31.9%)	15 (31.9%)	10 (21.3%)	7 (14.9%)	106 (100%)	2.8533	0.4148
Resíduos	0.975	0.857	-0.559	-1.334			
Defensores	86 (37.4%)	62 (26.9%)	48 (20.9%)	34 (14.8%)	405 (100%)	20.58	<0.001
Resíduos	4.057	0.191	-1.380	-2.992			
Meias	87 (32.1%)	94 (34.7%)	54 (19.9%)	36 (13.3%)	202 (100%)	25.271	<0.001
Resíduos	2.412	3.094	-1.857	-3.833			
Atacantes	53 (32.5%)	43 (26.4%)	38 (23.3%)	29 (17.8%)	163 (100%)	8.1451	0.043
Resíduos	2.692	-0.222	-0.856	-1.680			

*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

Diante dos resultados apresentados nas Tabelas 6 e 7, concluímos que é essencial avaliar a presença do Efeito da Idade Relativa (EIR) de forma segmentada. Nas Tabelas 8, 9 e 10, são fornecidos os resultados dos testes qui-quadrado de aderência para as combinações de década e posição dos jogadores.

Para os Defensores, podemos concluir que os nascidos nas décadas de 1990 ($p=0,003$) e 2000 ($p=0,001$) experimentaram a presença do Efeito da Idade Relativa em suas categorias de base, ao passo que os nascidos na década de 1980 ($p=0,785$) não apresentaram evidências desse efeito (Tabela 8).

Tabela 8. Testes qui-quadrado de aderência para os defensores - década de nascimento

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total	χ^2	P-valor
População BR*	25.6%	26.4%	24.8%	23.2%	100%		
Defensores de 1980	7 (19.4%)	10 (27.8%)	11 (30.6%)	8 (22.2%)	36 (100%)	1.0669	0.7851
Resíduos	-0.858	0.187	0.799	-0.124			
Defensores de 1990	50 (37.9%)	38 (28.8%)	24 (18.2%)	20 (15.1%)	142 (100%)	13.845	0.003
Resíduos	3.202	0.622	-1.760	-2.166			
Defensores de 2000	29 (46.8%)	14 (22.6%)	13 (21%)	6 (9.6%)	62 (100%)	16.26	0.001
Resíduos	3.797	-0.682	-0.698	-2.507			

*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

O mesmo padrão se repete para os Meias, em que observamos a presença do EIR entre 1990 ($p<0,001$) e 2000 ($p=0,003$), mas não para o grupo nascido na década de 1980 ($p=0,767$). Vale ressaltar que na década de 2000, há uma observação interessante, com mais jogadores nascidos no segundo trimestre do que no primeiro (Tabela 9).

Tabela 9. Testes qui-quadrado de aderência para os meias - década de nascimento

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total	χ^2	P-valor
População BR*	25.6%	26.4%	24.8%	23.2%	100%		
Meias de 1980	6 (22.2%)	9 (33.3%)	5 (18.5%)	7 (25.9%)	27 (100%)	1.1416	0.767
Resíduos	-0.413	0.817	-0.755	0.343			
Meias de 1990	59 (34.7%)	54 (31.7%)	36 (21.2%)	21 (12.3%)	190 (100%)	16.618	<0.001
Resíduos	2.687	1.588	-1.094	-3.324			
Meias de 2000	22 (29.7%)	31 (41.9%)	13 (17.6%)	8 (10.8%)	54 (100%)	13.594	0.003
Resíduos	3.797	-0.682	-0.698	-2.507			

*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

Quanto aos Atacantes, para a década de 1980 não houve significância no teste ($p=0,538$). Na década de 1990, embora o resultado do teste não seja significativo ($p=0,4504$), a tendência nos valores, de forma descritiva, sugere uma semelhança com o conceito do EIR, e a não significância pode ser atribuída ao tamanho amostral. Por fim, para a década de 2000, há evidência da presença significativa do Efeito da Idade Relativa ($p=0,023$) (Tabela 10).

Tabela 10. Testes qui-quadrado de aderência para os atacantes - por década de nascimento

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total	χ^2	P-valor
População BR*	25.6%	26.4%	24.8%	23.2%	100%		
Atacantes de 1980	5 (20.8%)	4 (16.7%)	8 (33.3%)	7 (29.2%)	24 (100%)	2.1695	0.538
Resíduos	-0.545	-1.081	0.968	0.705			
Atacantes de 1990	23 (30.3%)	20 (26.3%)	19 (25%)	14 (18.4%)	77 (100%)	2.6409	0.4504
Resíduos	1.466	-0.218	-0.154	-1.134			
Atacantes de 2000	25 (39.7%)	19 (30.1%)	11 (17.5%)	8 (12.7%)	63 (100%)	9.4491	0.023
Resíduos	2.539	0.676	-1.349	-1.958			

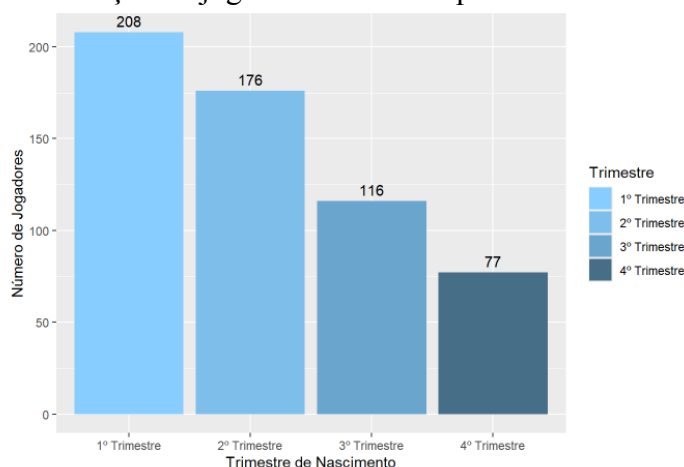
*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

4.2 Grupos de jogadores com a presença do Efeito da Idade Relativa

Com base nas informações apresentadas na seção 4.1, observamos que não dispomos de evidências que sustentem a consideração da presença do efeito da idade relativa (EIR) nos jogadores da Série A de 2023 nascidos na década de 1980. Além disso, tanto nesta amostra quanto na revisão da literatura, não encontramos confirmações desse efeito em goleiros. Diante desse cenário, optamos por excluir esses jogadores da análise que será apresentada nesta seção 4.2.

Na Figura 3, observamos a distribuição dos 577 jogadores que identificamos como tendo vivenciado o efeito da idade relativa durante suas categorias de base. Esses jogadores incluem defensores, meias e atacantes nascidos nas décadas de 1990 e 2000.

Figura 3. Distribuição de jogadores com EIR por trimestre de nascimento



Adicionalmente, na Tabela 11, apresentamos os resultados do teste qui-quadrado de aderência para o grupo de jogadores com EIR em comparação com a população de referência. O p-valor < 0,001 sustenta a hipótese de que esses jogadores experienciaram o EIR em suas categorias de base.

Tabela 11. Teste qui-quadrado de aderência para os jogadores com EIR

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total	χ^2	P-valor
População BR*	25.6%	26.4%	24.8%	23.2%	100%		
Jogadores do BR 23	208 (36%)	176 (30.5%)	77 (20.1%)	77 (13.4%)	577 (100%)	46.851	<0.001
Resíduos	5.688	2.235	-2.612	-5.559			

*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

4.3 Grupos de jogadores para análise quantitativa

Após a identificação dos jogadores que foram impactados pelo Efeito da Idade Relativa (EIR), procedemos para a fase de análise das variáveis quantitativas. Com o intuito de mitigar possíveis efeitos de confundimento nas análises, excluímos aqueles jogadores que não estiveram em suas equipes desde o início do campeonato. Essa exclusão foi realizada para evitar que medidas fossem influenciadas por fatores externos à pesquisa, uma vez que esses jogadores não tiveram a oportunidade de integrar o elenco ao longo das 38 rodadas do campeonato. Nesse contexto, foram excluídos 137 jogadores, conforme informações disponíveis no site *TransferMarkt*⁴, que tiveram movimentações em suas carreiras durante a janela de contratações de julho de 2023.

É pertinente observar, por meio da Tabela 12, que os 440 jogadores que permaneceram na análise apresentam uma distribuição de nascimento por trimestre muito

⁴https://www.transfermarkt.com.br/campeonato-brasileiro-serie-a/transfers/wettbewerb/BRA1/saison_id/2023/s_w/s

similar à do grupo de jogadores identificados na seção 4.2 (Tabela 11). Essa consistência reforça a robustez da amostra utilizada na análise das variáveis quantitativas, contribuindo para a validade e confiabilidade dos resultados obtidos.

Tabela 12. Distribuição dos jogadores que estiveram em todo o campeonato

	1º Trimestre	2º Trimestre	3º Trimestre	4º Trimestre	Total
População BR*	25.6%	26.4%	24.8%	23.2%	100%
Jogadores do BR 23	160 (36.3%)	139 (31.6%)	83 (18.9%)	58 (13.2%)	440 (100%)

*Fonte: Datasus (<http://tabnet.datasus.gov.br/>)

Com base nisso, procedemos à análise de 5 variáveis quantitativas, a saber: Participação no campeonato (número de jogos disputados / 38 possíveis); Titularidade (número de jogos em que começou como titular / número de jogos disputados); Média de Minutos Por Jogo disputado (número de minutos jogados no campeonato / número de jogos disputados); Minutos Jogados (número de minutos jogados no campeonato); Nota *SofaScore* (Nota de desempenho atribuída pelo site *SofaScore*).

Na Tabela 13, apresentamos um resumo dessas variáveis, com a média e o respectivo intervalo de confiança de 95% para cada trimestre de nascimento. É interessante notar que o 4º Trimestre se destaca por apresentar os valores mais elevados nas 5 variáveis, embora as diferenças sejam pequenas. Essa observação sugere uma tendência que merece uma análise mais aprofundada, pois pode indicar uma possível influência do Efeito da Idade Relativa nas variáveis em questão.

Tabela 13. Médias das variáveis quantitativas por Trimestre

Trimestre	Participação	Titularidade	Minutos Por Jogo	Minutos Jogados	Nota SofaScore
1º Trimestre	46.15% (41.43;50.87)	56.26% (51.16;61.35)	54.29 (50.83;57.75)	1074.29 (935.69;1212.88)	6.727 (6.694;6.760)
2º Trimestre	46.68% (41.37;51.99)	53.15% (47.16;59.14)	51.21 (47.18;55.23)	1116.99 (955.31;1278.68)	6.743 (6.702;6.783)
3º Trimestre	48.03% (41.88;54.18)	52.34% (45.27;59.41)	50.80 (45.79;55.81)	1101.13 (916.51;1285.76)	6.752 (6.709;6.795)
4º Trimestre	49.36% (41.59;57.13)	58.89% (49.88;67.91)	56.44 (49.99;62.88)	1213.83 (965.67;1461.98)	6.779 (6.727;6.831)

Contribuindo com a análise das variáveis em questão, entre as Figuras 4 e 8 apresentamos *boxplots* das 5 variáveis, a fim de observarmos como se comporta a distribuição dessas variáveis. Em suma, as distribuições apresentam amplitudes similares, com exceção da “Nota *SofaScore*” (Figura 8), onde as distribuições do 1º e 2º Trimestres possuem uma amplitude maior que a do 4º Trimestre. Podemos ver que esses trimestres possuem algumas notas que destoam para baixo, as quais o site *SofaScore* (2019)⁵ define da

⁵ Acessado em 17/12/2023 em (<https://www.sofascore.com/news/pontuacoes-estatisticas-sofascore/>)

seguinte forma: “Em geral, os jogadores raramente recebem pontuações abaixo de 6, e essa pontuação normalmente é um indicador de erro maior como um gol contra, pênalti concedido, cartão vermelho ou simplesmente uma soma de erros.”

Figura 4. Boxplots para a proporção média de participação no campeonato

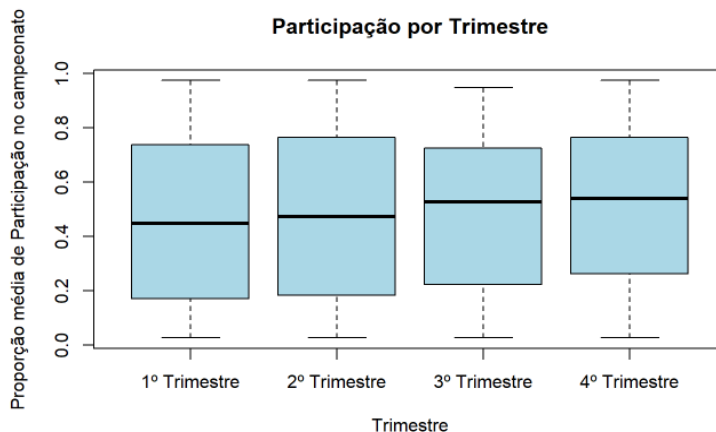


Figura 5. Boxplots para a proporção média de titularidade no campeonato

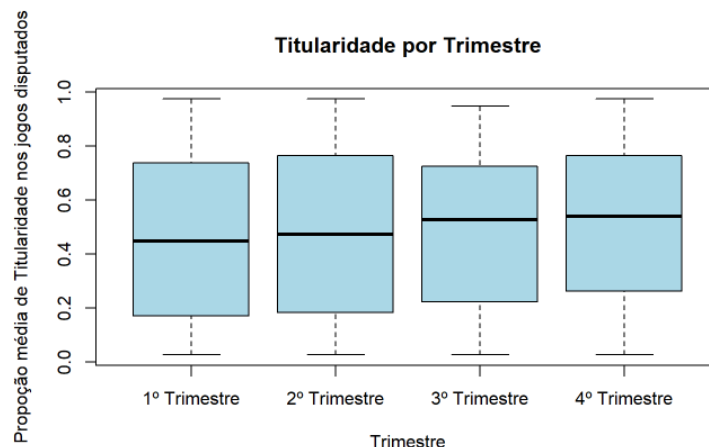


Figura 6. Boxplots para a média de minutos jogados no campeonato

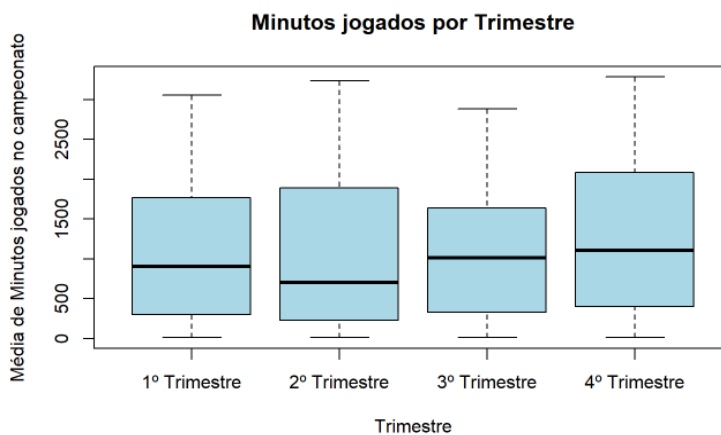
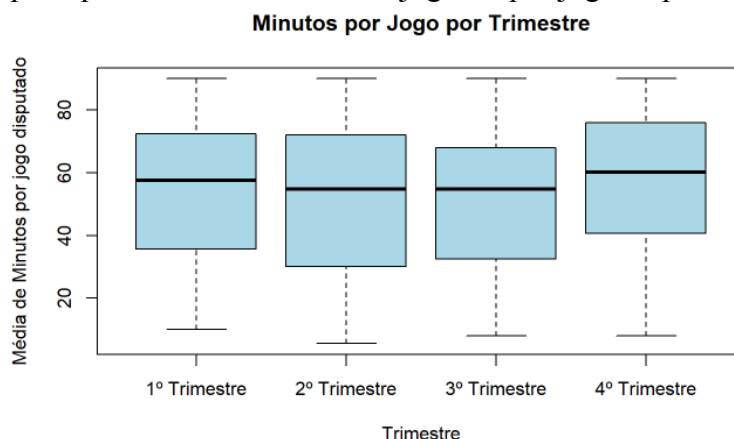
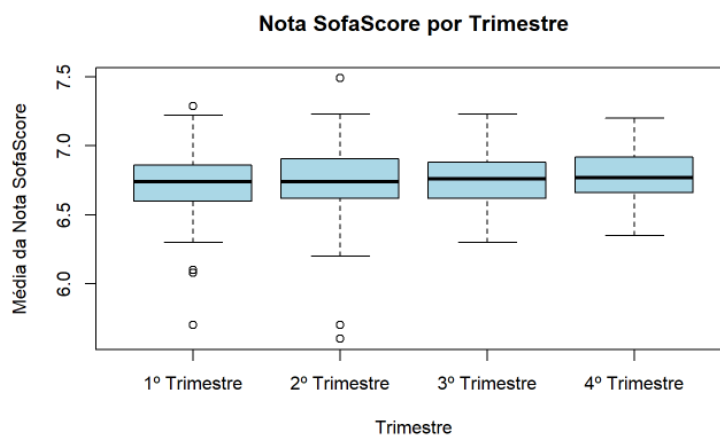


Figura 7. Boxplots para a média de minutos jogados por jogo disputado no campeonato

Figura 8. Boxplots para a média da Nota *SofaScore* no campeonato

Após extensivas análises exploratórias das variáveis, não foram identificados padrões significativos para este estudo quando consideramos apenas a métrica de desempenho e o trimestre de nascimento. No entanto, quando subdividimos os grupos por posição dos jogadores, surgem padrões distintos nas medidas para cada posição. Notavelmente, no caso dos Defensores, destaca-se uma tendência clara para medidas mais elevadas nos jogadores nascidos no 4º trimestre (Tabela 14).

Tabela 14. Médias das variáveis quantitativas por Trimestre - Defensores

Trimestre	Participação	Titularidade	Minutos Por Jogo	Minutos Jogados	Nota SofaScore	*
1º Trimestre	38.46% (30.95;45.96)	68.30% (60.34;76.27)	50.89 (45.66;56.12)	1184.90 (975.26;1394.54)	6.70 (6.63;6.76)	a-
2º Trimestre	44.98% (35.48;54.48)	69.35% (59.84;78.85)	49.21 (43.72;54.71)	1155.06 (927.63;1382.49)	6.75 (6.67;6.84)	ab
3º Trimestre	48.52% (38.80;58.24)	63.06% (52.05;74.07)	47.87 (40.27;55.47)	1111.06 (814.91;1407.21)	6.77 (6.70;6.83)	ab
4º Trimestre	48.44% (34.47;62.42)	77.61% (63.70;91.53)	56.53 (48.35;64.70)	1278.44 (892.09;1664.78)	6.86 (6.77;6.96)	-b

* Resultado para as comparações múltiplas Tukey

Especificamente, ao conduzir uma análise de variância (ANOVA) para a variável "Nota SofaScore", identificou-se significância estatística com um p-valor = 0,0472. Isso sugere que existe uma diferença significativa na Nota de Desempenho entre os Trimestres de Nascimento para os Defensores. Além disso, por meio do teste de comparações múltiplas Tukey, verificou-se uma diferença significativa com um p-valor = 0,029 entre o 4º e o 1º Trimestre.

Esses resultados destacam uma variação significativa no desempenho dos Defensores em relação ao trimestre de nascimento, especialmente evidenciada pela métrica "Nota SofaScore". Essas descobertas específicas podem contribuir significativamente para o entendimento mais profundo das influências do Efeito da Idade Relativa em diferentes posições no contexto do futebol.

Na Tabela 15 e 16, notamos o comportamento das variáveis para Meias e Atacantes. No caso dos Meias, não há um padrão tão pronunciado, entretanto, três variáveis ainda indicam medidas mais elevadas para o 4º Trimestre, que são Titularidade, Minutos Por Jogo e Minutos Jogados. Quanto aos Atacantes, algo curioso ocorre, uma vez que os valores se invertem nessa posição do campo. As medidas mais elevadas concentram-se no 1º Trimestre, ao contrário do que observamos para os Defensores.

Tabela 15. Médias das variáveis quantitativas por Trimestre - Meias

Trimestre	Participação	Titularidade	Minutos Por Jogo	Minutos Jogados	Nota SofaScore
1º Trimestre	53.90% (46.84;60.97)	54.60% (46.94;62.26)	65.28 (60.17;70.39)	1054.44 (815.99;1292.88)	6.76 (6.72;6.81)
2º Trimestre	50.75% (42.96;58.53)	51.66% (42.93;60.39)	64.18 (57.66;70.71)	1287.86 (972.35;1603.37)	6.75 (6.69;6.81)
3º Trimestre	50.45% (40.65;60.25)	50.73% (39.80;61.66)	60.55 (52.70;68.40)	1256.00 (945.68;1566.32)	6.76 (6.69;6.84)
4º Trimestre	52.17% (39.52;64.83)	59.84% (46.97;72.70)	70.25 (59.57;80.93)	1466.00 (997.99;1934.01)	6.76 (6.68;6.83)

Tabela 16. Médias das variáveis quantitativas por Trimestre - Atacantes

Trimestre	Participação	Titularidade	Minutos Por Jogo	Minutos Jogados	Nota SofaScore
1º Trimestre	46.05% (34.99;57.11)	38.35% (27.97;48.74)	41.21 (34.66;47.76)	917.97 (613.87;1222.07)	6.72 (6.65;6.78)
2º Trimestre	39.84% (27.72;51.95)	32.58% (20.65;44.51)	36.57 (28.23;44.91)	775.69 (438.44;1112.94)	6.70 (6.62;6.78)
3º Trimestre	41.78% (25.46;58.10)	34.44% (18.29;50.60)	37.71 (27.32;48.10)	769.69 (383.99;1155.38)	6.70 (6.60;6.80)
4º Trimestre	45.95% (28.73;63.17)	25.55% (13.70;37.40)	32.91 (23.68;42.13)	672.77 (302.26;1043.28)	6.67 (6.58;6.76)

Tanto para os Meias quanto para os Atacantes, não observou-se significância estatística ao conduzir os modelos ANOVA. Além disso, cabe ressaltar que todos os pressupostos foram atendidos ao proceder essas análises.

5 CONCLUSÕES

Ao término desta pesquisa, é possível derivar conclusões relevantes a partir da análise do Efeito da Idade Relativa (EIR) sobre os jogadores de futebol na Série A do Campeonato Brasileiro de 2023. As conclusões são delineadas a partir dos resultados obtidos e das análises realizadas ao longo do trabalho.

5.1 Efeito da Idade Relativa: Evidências e Variações

Os resultados deste estudo corroboram com os estudos anteriores, tais como o de Massa (2017) e Marques, Pinheiro e Coswig (2018) e confirmam a presença do Efeito da Idade Relativa na Série A do Campeonato Brasileiro de 2023. Essa constatação é respaldada por testes qui-quadrado de aderência, que revelaram uma disparidade estatisticamente significativa entre a distribuição de nascimentos na população de referência e a composição da amostra.

Além disso, ao considerar a variabilidade temporal, observou-se que a presença do EIR não é uniforme ao longo das décadas. Enquanto nas décadas de 1990 e 2000 há evidências significativas desse efeito, na década de 1980 não se observou a mesma influência estatisticamente relevante. Essa diferenciação destaca a importância de levar em conta a evolução do contexto esportivo ao investigar o impacto do EIR em diferentes épocas do futebol.

É importante destacar que esse achado pode ter sido influenciado pelo tamanho da amostra dos jogadores nascidos na década de 1980 ($n = 103$). Além disso, essa variação pode ocorrer devido à evolução do futebol e, conseqüentemente, dos processos de escolha dos jogadores nas categorias de base, à medida que o futebol se torna mais competitivo e decidido em detalhes.

Ainda, uma contribuição marcante desta pesquisa é a observação do impacto diferenciado do Efeito da Idade Relativa em diferentes posições do campo. Os testes qui-quadrado de aderência indicaram a presença do EIR entre Defensores, Meias e Atacantes, mas não entre goleiros. Além disso, as variáveis quantitativas se comportaram de maneiras distintas entre as posições. Esse achado reforça a ideia de que o EIR não se manifesta

uniformemente em todas as posições, ressaltando a necessidade de abordagens segmentadas ao avaliar suas variações.

5.2 Tendências de Desempenho

Ao analisar as variáveis quantitativas relacionadas ao desempenho dos jogadores, identificamos uma tendência notável: jogadores nascidos no 4º Trimestre apresentaram, em média, valores mais elevados em diversas métricas. A análise de variância (ANOVA) e o teste de comparações múltiplas Tukey confirmaram a significância estatística dessa diferença entre os defensores, na variável "Nota SofaScore".

Essas descobertas têm implicações práticas significativas, sugerindo que jogadores nascidos no último trimestre podem ter um desempenho estatisticamente superior em comparação com seus colegas nascidos em outros trimestres. Isso pode influenciar decisões de treinadores, gestores e olheiros no contexto do futebol, destacando a importância de considerar o Efeito da Idade Relativa ao avaliar e desenvolver talentos.

Essa constatação está em consonância com o que Neto et al (2020) discutem em sua publicação. Na época, analisando apenas a quantidade de minutos jogados no campeonato, o 4º Trimestre já apresentava vantagem. Segundo os autores, isso pode se dever ao fato de que, para superar as desvantagens de ser do 4ª Trimestre do ano nas categorias de base, os jogadores precisam ter um desempenho excepcional. Assim, teríamos a hipótese de que apenas os melhores do 4º Trimestre chegam ao futebol profissional, fazendo com que as métricas de desempenho tenham um comportamento mais eficiente. Agora, com a inclusão de mais variáveis na análise, observamos uma tendência semelhante para os Defensores. Entretanto, é importante destacar que percebemos que esse efeito não se manifesta de maneira uniforme entre as diferentes posições. Portanto, é imprescindível que, em futuros trabalhos, isso seja levado em consideração.

5.3 Limitações e Perspectivas para Pesquisas Futuras

Como em qualquer pesquisa, é essencial reconhecer suas limitações. Neste estudo, a análise concentrou-se na Série A do Campeonato Brasileiro, e extensões para outras ligas e categorias poderiam enriquecer ainda mais a compreensão do Efeito da Idade Relativa no futebol. Perspectivas futuras também podem explorar fatores adicionais que podem modular a

influência do EIR, como características socioeconômicas e experiências de treinamento. Adicionalmente, considerar dinâmicas específicas das categorias de base e processos de seleção pode oferecer *insights* mais detalhados sobre a formação e desenvolvimento de jogadores.

5.4 Contribuições e Relevância

Este trabalho pode contribuir para a literatura ao evidenciar a presença do Efeito da Idade Relativa no contexto do futebol brasileiro, fornecendo *insights* sobre suas variações temporais e impactos diferenciados nas posições de campo. As implicações práticas destacam a importância de uma abordagem informada ao avaliar e gerenciar talentos no futebol.

Essas conclusões consolidam os resultados e a relevância do estudo, proporcionando uma visão abrangente do impacto do Efeito da Idade Relativa no cenário do futebol brasileiro em 2023. Além disso, a partir deste trabalho e dos já existentes na literatura, tais como Neto et Al (2020), Massa (2017) e Williams (2010), podemos afirmar que o futebol precisa encontrar formas de diminuir o Efeito da Idade Relativa nos processos de escolha das categorias de base. Se esse problema persistir, corremos o risco de perder talentos que poderiam se tornar bons profissionais no mundo do futebol.

REFERÊNCIAS

- AGRESTI, A. (2007). An introduction to categorical data analysis. John Wiley & Sons.
- ANDERSON, C., & SALLY, D. (2013). *The Numbers Game: Why Everything You Know About Soccer Is Wrong*. Penguin UK.
- BARNSLEY, R. H., THOMPSON, A. H., & BARNSLEY, P. E. (1985). Hockey success and birthdate: The relative age effect. *Canadian Association for Health, Physical Education, and Recreation Journal*, 51(1), 23-28.
- CRUZ, J. (2012). A formação de atletas de futebol: um estudo na categoria sub-15 do Cruzeiro Esporte Clube, Belo Horizonte-MG. *Revista Brasileira de Futsal e Futebol*. São Paulo. Vol. 4. Núm. 13.
- DELORME, N., & RASPAUD, M. (2009). The relative age effect in elite sports: The French case. *Research Quarterly for Exercise and Sport*, 80(2), 336-344.
- DELORME N., & RASPAUD M.(2009). Is there an influence of relative age on participation in non-physical sports activities? The example of shooting sports. *J Sports Sci.*; 27(10): 1035-1042.
- DICKINSON, D. J.; LARSON(1963), J. D. The effects of chronological age in months on school achievement. *The Journal of Educational Research*. Vol. 56. Núm. 9. p. 492-493. 1963.
- GUILY(1965) , M. Mois de naissance et réussite scolaire. *Enfance*. Vol. 18. Núm. 4. p. 491-503. 1965.
- HIROSE, N. (2009). Relationships among birth-month distribution, skeletal age and anthropometric characteristics in adolescent elite soccer players. *Journal of Sports Sciences*, 27(11), 1159-1166.
- HOSMER, D. W., LEMESHOW, S., & STURDIVANT, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- MALINA, R. M., MORANO, P. J., BARRON, M., MILLER, S. J., & CUMMING, S. P. (2009). Influence of relative age on self-regulation in male youth soccer. *Journal of Sports Sciences*, 27(10), 1073-1079.
- MARQUES, E. A., PINHEIRO, A. R., & COSWIG, V. S. (2018). Efeito da idade relativa sobre a seleção de atletas para as categorias de base de um clube de futebol. *Rev Bras Ciênc Esporte*. 41(2):157-162.
- MASSA, M. (2017). Efeito da idade relativa e maturação biológica: implicações para seleção de jovens jogadores de futebol. Tese de Livre Docência, Universidade de São Paulo, São Paulo.
- MASSA, M., et al. (2020). O efeito da idade relativa nas delegações olímpicas brasileiras. *R. bras. Ci. e Mov* 2020;28(4):42-51.
- MOREIRA, A., et al. (2013). Influence of sport and non-sporting activities on the spatial-temporal parameters of the gait cycle in children and adolescents. *Gait & Posture*, 38(4), 864-869.
- MUSCH, J., & GRONDIN, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review*, 21(2), 147-167.
- NETO, E., et al. (2020). Influência da idade relativa na participação de jogadores de futebol na Série A do campeonato brasileiro. *Rev Bras Futebol* 2020; v. 13, n. 3, 41 – 53.

PEARSON, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine, Series 5*, 50(302), 157-175.

SHAMAH, M. (2021). Análise de desempenho no futebol: A prática do analista de desempenho nas categorias de base dos clubes brasileiros da Série A. Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre.

WILLIAMS, J. H. (2010). Relative age effect in youth soccer: Analysis of the FIFA U17 World Cup competition. *Scandinavian Journal of Medicine & Science in Sports*, 20(3), 502-508.