



**Universidade de Brasília  
Departamento de Estatística**

**Análise Multivariada por fatores determinantes do desempenho  
de jogadores do Brasileirão Série A - 2022**

**Francisco Iago dos Reis Ferreira**

Projeto apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília  
2023**

**Francisco Iago dos Reis Ferreira**

**Análise Multivariada por fatores determinantes do desempenho  
de jogadores do Brasileirão Série A - 2022**

Orientador: Prof. Gladston Luiz da Silva

Projeto apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília  
2023**

---

# Agradecimentos

Foi uma longa jornada até aqui, agradeço a todas as pessoas que contribuíram para a realização desse trabalho.

À minha mãe, a Sra. Francislene Maria dos Reis, por ter batalhado para que minha irmã e eu pudessemos ter a melhor educação possível. Todo o seu esforço não foi em vão.

À minha companheira, Evellyn Cristina Lopes Guerbe, cujo o amor e apoio incondicional foram cruciais para superar os momentos de indecisão.

Aos vários amigos que fiz durante toda a graduação, seja no curso de estatística, seja na Casa do Estudante, tenho muita consideração por cada um de vocês.

À meu orientados, Prof. Gladston Luis da Silva, pela empolgação em orientar esse trabalho, sua orientação e conselhos foram fundamentais para aprimorar minha pesquisa e direcionar meus esforços na direção certa.

E por último, à todos os professores do Departamento de Estatística e à Universidade de Brasília, por proporcionarem o ambiente de aprendizado e os recursos necessários para a concretização deste projeto.

---

# Resumo

O Campeonato Brasileiro, também conhecido como Brasileirão, é a principal competição de futebol profissional no Brasil. sendo disputado todos os anos por clubes que representam diversos estados de todo o país. Adotado em 2003, o formato de pontos corridos é disputado por 20 clubes que se enfrentam em jogos de ida e volta durante 38 rodadas, a equipe que fizer a maior pontuação se sagra como a campeã daquela edição. Esse trabalho teve como objetivo empregar técnicas de Análise Multivariada com o intuito de identificar os principais fatores que influenciam o desempenho dos clubes participantes do Campeonato a partir de dados relativos aos jogadores que integram os times na competição disputada em 2022.

**Palavras-chaves:** Brasileirão, Desempenho, Análise Fatorial, Análise de Componentes Principais.

---

# Abstract

The Brazilian Championship, also known as the Brasileirão, is the main professional soccer competition in Brazil and is contested every year by clubs representing various states across the country. Adopted in 2003, the straight points format is played by 20 clubs who face each other in back-to-back matches over 38 rounds, with the team that scores the highest being crowned champion of that edition. The aim of this study was to use Multivariate Analysis techniques to identify the main factors influencing the performance of the clubs taking part in the Championship, using data on the players who make up the teams in the competition in 2022.

**Keywords:** Brasileirão, Performance, Factor Analysis, Principal Component Analysis.

## Lista de Tabelas

1	Descrição das variáveis dos defensores - Geral . . . . .	23
2	Análise de Componentes Principais dos defensores - Geral . . . . .	25
3	Autovetores para CP1 e CP2 dos defensores . . . . .	26
4	Análise de Componentes Principais dos defensores - G6 e Z4 . . . . .	29
5	Autovetores relacionados a G6 e Z4 - Defensores . . . . .	29
6	Descrição das variáveis dos meio campistas - Geral . . . . .	31
7	Análise de Componentes Principais dos meio campistas - Geral . . . . .	32
8	Autovetores para CP1 e CP2 dos meio campistas . . . . .	33
9	Análise de Componentes Principais dos meio campistas - G6 e Z4 . . . . .	36
10	Autovetores relacionados a G6 e Z4 - Meio campistas . . . . .	36
11	Descrição das variáveis dos atacantes - Geral . . . . .	38
12	Análise de Componentes Principais dos atacantes - Geral . . . . .	39
13	Autovetores para CP1 e CP2 dos atacantes . . . . .	40
14	Análise de Componentes Principais dos atacantes - G6 e Z4 . . . . .	43
15	Autovetores relacionados a G6 e Z4 - Atacantes . . . . .	43
16	TOP 10 Melhores defensores (CP1) - G6 . . . . .	45
17	TOP 10 Melhores defensores (CP1) - Z4 . . . . .	46
18	TOP 10 Melhores defensores (CP2) - G6 . . . . .	46
19	TOP 10 Melhores defensores (CP2) - Z4 . . . . .	47
20	TOP 10 Melhores meio campistas (CP1) - G6 . . . . .	47
21	TOP 10 Melhores meio campistas (CP1) - Z4 . . . . .	48
22	TOP 10 Melhores meio campistas (CP2) - G6 . . . . .	48
23	TOP 10 Melhores meio campistas (CP2) - Z4 . . . . .	49
24	TOP 10 Melhores atacantes (CP1) - G6 . . . . .	49
25	TOP 10 Melhores atacantes (CP1) - Z4 . . . . .	50
26	TOP 10 Melhores atacantes (CP2) - G6 . . . . .	50

27	TOP 10 Melhores atacantes (CP2) - Z4 . . . . .	51
----	--	----

## **Lista de Figuras**

1	Metodologia CRISP DM . . . . .	11
2	Tabela após a 38 <sup>a</sup> Rodada . . . . .	20
3	Pagina inicial - Campeonato Brasileiro 2022 . . . . .	21
4	Classificação dos clubes . . . . .	21
5	Marcação correta dos dados . . . . .	22
6	Correlação entre as variáveis - Defensores . . . . .	24
7	Variáveis dos defensores - Geral . . . . .	27
8	Correlações dos defensores - G6 . . . . .	28
9	Correlações dos defensores - Z4 . . . . .	28
10	Variáveis dos defensores - G6 e Z4 . . . . .	30
11	Correlação entre as variáveis - Meio Campistas . . . . .	32
12	Variáveis dos meio campistas - Geral . . . . .	34
13	Correlações dos meio campistas - G6 . . . . .	35
14	Correlações dos meio campistas - Z4 . . . . .	35
15	Variáveis dos meio campistas - G6 e Z4 . . . . .	37
16	Correlação entre as variáveis - Atacantes . . . . .	39
17	Variáveis dos atacantes - Geral . . . . .	41
18	Correlações dos atacantes - G6 . . . . .	42
19	Correlações dos atacantes - Z4 . . . . .	42
20	Variáveis dos atacantes - G6 e Z4 . . . . .	44

## Sumário

<b>Capítulo 1 - Introdução</b> . . . . .	8
1.1 Objetivo Geral . . . . .	10
1.1.1 Objetivos Específicos . . . . .	10
1.2 Metodologia. . . . .	11
<b>Capítulo 2 - Método Estatístico</b> . . . . .	13
2.1 Análise Multivariada. . . . .	13
2.2 Análise de Componentes Principais . . . . .	13
2.3 Trabalhos Relacionados . . . . .	17
<b>Capítulo 3 - Resultados</b> . . . . .	19
3.1 Campeonato Brasileiro de Futebol de 2022 - Série A . . . . .	19
3.2 Coleta dos dados . . . . .	21
3.2.1 Tipos de Variáveis . . . . .	22
3.2.2 Tratamento e divisão dos dados . . . . .	22
3.3 Defensores. . . . .	23
3.3.1 Aplicação de Componentes Principais . . . . .	24
3.3.2 Autovetores . . . . .	26
3.3.3 Comparação entre os Grupos . . . . .	27
3.4 Meio Campistas . . . . .	30
3.4.1 Análise Componentes Principais . . . . .	32
3.4.2 Comparação entre os Grupos . . . . .	34
3.5 Atacantes . . . . .	37
3.5.1 Análise dos Componentes Principais . . . . .	39
3.5.2 Comparação entre os grupos . . . . .	41
3.6 Criação dos Escores . . . . .	44
3.6.1 Defensores . . . . .	45
3.6.2 Meio Campistas . . . . .	47

3.6.3 Atacantes . . . . .	49
<b>Conclusões . . . . .</b>	<b>52</b>

## Capítulo 1 - Introdução

No ano de 2002, Billy Beane, diretor geral da equipe de beisebol *Oakland Athletics*, desafiou tudo que se acreditava sobre o esporte competitivo até então, ao adotar técnicas estatísticas avançadas para avaliar o desempenho de jogadores. Contrapondo-se aos critérios que eram utilizados pelos olheiros à época, como capacidade de correr, de arremessar, receber a bola, de rebater e rebater com potência, Billy e Paul dePodesta (até então um jovem executivo do clube) foram capazes de concluir que nem todas as características importantes num jogador de beisebol têm a mesma importância. A velocidade na corrida, habilidade na defesa e força bruta tendiam a ser características demasiadamente superestimadas.

Eles foram capazes de mostrar que características como porcentagem de rebatidas, porcentagem de bases alcançadas, habilidade de controle da base de *strike*, entre outras métricas estatísticas, eram indicadores mais confiáveis do desempenho de um jogador do que as avaliações tradicionais. Essa abordagem revelou informações valiosas sobre a eficácia de um jogador, permitindo que talentos muitas vezes subvalorizados por outras equipes fossem identificados.

Com um orçamento limitado, sendo o time com a terceira menor folha salarial da liga na época, mas com uma abordagem estatística autêntica, o time do Oakland Athletics quebrou o recorde de vitórias consecutivas de uma temporada (LEWIS, 2004). O impacto dessa abordagem abriu caminho para uma nova era no esporte, extrapolou o beisebol e inspirou diversas modalidades. A utilização de técnicas estatísticas promoveu uma mudança na maneira como o desempenho dos jogadores seria avaliado e compreendido.

No âmbito do futebol, é imperativo considerar o avanço relativo à informação contida nos dados. De acordo com Anderson e Sally (2013), os números têm o poder de desafiar conceitos pré-estabelecidos e dismantelar crenças antigas no esporte. Eles oferecem um entendimento do jogo como nunca visto antes, possibilitando uma visão contraposta aos pressupostos tradicionais.

Um exemplo de como os dados podem influenciar na tomada de decisão é a renovação de contrato feita pelo jogador do Manchester City, Kevin de Bruyne<sup>1</sup>. Na época, o jogador conseguiu uma extensão salarial na casa dos 80 milhões de euros, utilizando apenas análise de dados ao invés de um agente<sup>2</sup>.

---

<sup>1</sup><https://www.transfermarkt.com.br/kevin-de-bruyne/profil/spieler/88755>, acessado em 15/04/2023

<sup>2</sup><https://www.uol.com.br/esporte/futebol/ultimas-noticias/2021/04/08/big-data-e-sem-agente-como-de-bruyne-renovou-com-o-city.htm>, acessado em 15/04/2023

No Brasil, o Brasileirão Série A é uma competição que movimenta milhões de reais, instiga torcedores, investidores e apostadores todos os anos. Com o crescente avanço da tecnologia, o uso de técnicas estatísticas tem sido de imensa importância para se avaliar e entender a performance dos atletas durante a competição. No cenário nacional, por exemplo, times como o Flamengo<sup>3</sup> e Palmeiras<sup>4</sup> utilizam análise de desempenho para tomada de decisões administrativas e esportivas.

O presente trabalho visa analisar, por meio de técnicas de Análise Multivariada, o desempenho de atletas da primeira divisão do Campeonato Brasileiro Série A 2022, pois permite uma avaliação mais precisa e completa, além de identificar padrões e tendências que não seriam encontrados por meio de análises mais simples. Por meio dessa análise, espera-se identificar as características que mais contribuíram para o desempenho dos jogadores durante o campeonato.

Tal abordagem aplicada ao contexto do campeonato brasileiro pode ser útil para os treinadores, apostadores, analistas de desempenho, investidores e patrocinadores, além de entusiastas por futebol, que poderão tomar melhores decisões no âmbito esportivo utilizando-se de informações munidas de análise estatística.

---

<sup>3</sup><https://encurtador.com.br/rALOV>, acessado em 10/04/2023

<sup>4</sup><https://encurtador.com.br/doyMZ>, acessado em 10/04/2023

## 1.1 Objetivo Geral

Este trabalho teve como objetivo empregar técnicas de Análise Multivariada com o intuito de identificar os principais fatores que influenciam o desempenho dos clubes participantes do Campeonato Brasileiro da Série A em 2022. Inicialmente, a análise foi conduzida para os 20 clubes, seguida de uma comparação entre o desempenho dos clubes que alcançaram a fase de grupos da Libertadores e aqueles que foram rebaixados à segunda divisão. Foram levadas em consideração as posições de linha dos jogadores dessas equipes.

### 1.1.1 Objetivos Específicos

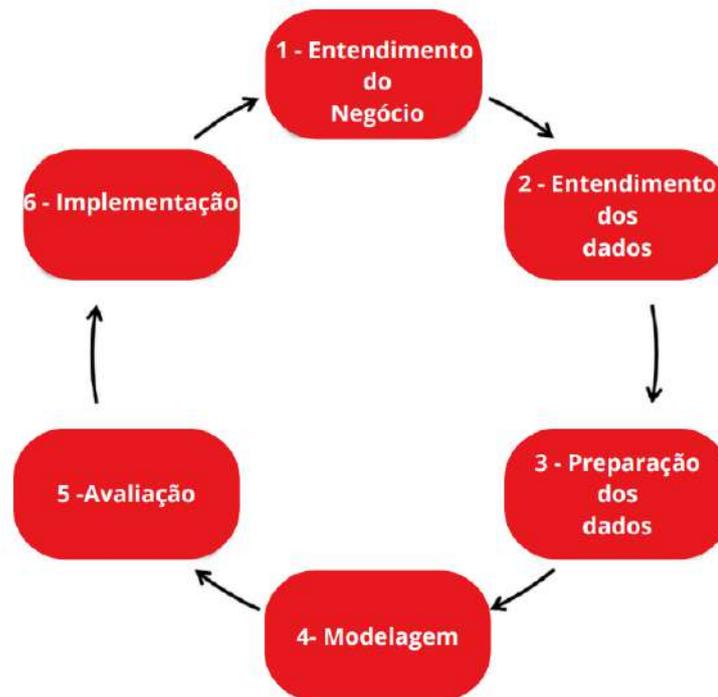
Para alcançar o objetivo geral, será necessário realizar os seguintes passos:

- Coletar e preparar os dados de desempenho dos jogadores que participaram do campeonato;
- Realizar uma análise exploratória dos dados para identificar possíveis valores ausentes e inconsistências;
- Aplicar técnicas de redução de dimensionalidade para selecionar as variáveis mais relevantes;
- Validar os resultados obtidos;
- Interpretar os resultados obtidos.

## 1.2 Metodologia

A elaboração deste trabalho baseou-se na metodologia CRISP-DM, visando organizar o trabalho de maneira eficaz e padronizada<sup>5</sup>.

Figura 1: Metodologia CRISP DM



Fonte: Adaptada pelo autor (*Smart Vision*)

A partir da Figura 1, podemos organizar o trabalho da seguinte forma:

- **Fase 1 - Entendimento do problema:** Na etapa inicial será feita a coleta e organização dos dados.
- **Fase 2 - Entendimento dos dados:** Nessa etapa, os dados ausentes e inconsistentes serão devidamente identificados e tratados.
- **Fase 3 - Preparação dos dados:** Após a etapa anterior, os dados serão preparados para a análise, isso inclui a padronização dos dados e a seleção das variáveis mais relevantes.
- **Fase 4 - Modelagem:** Serão utilizados métodos de Análise Multivariada. Além de realizarmos testes para validação dos resultados obtidos.

<sup>5</sup><https://www.sv-europe.com/crisp-dm-methodology/>, acessado em 10/09/2023

- **Fase 5 - Avaliação:** Os resultados serão avaliados em relação aos objetivos deste trabalho. Caso necessário, serão realizadas análises adicionais para obter resultados mais pertinentes.
- **Fase 6 - Conclusão:** Os resultados obtidos serão discutidos e validados.

## Capítulo 2 - Método Estatístico

Neste capítulo serão descritos os métodos de Análise Estatística que foram empregados na pesquisa. Tais métodos foram selecionados devido à sua relevância dentro do contexto dos dados coletados e contribuirão para alcançar os objetivos descritos na seção anterior.

### 2.1 Análise Multivariada

A Análise Multivariada é um conjunto de métodos estatísticos usados quando múltiplas variáveis são medidas simultaneamente em cada elemento amostral. Geralmente, são variáveis que estão correlacionadas entre si, e quanto maior o número, mais complexa se torna a análise. O objetivo é simplificar a interpretação do fenômeno estudado. Ela pode ser dividida em duas categorias principais: técnicas exploratórias que buscam simplificar a estrutura de variabilidade dos dados, sintetizando as variáveis, e técnicas de inferência. Um dos métodos comuns de simplificação da estrutura de variabilidade é a redução de dimensionalidade, que visa diminuir o número de variáveis originais mantendo as informações essenciais para a análise.

### 2.2 Análise de Componentes Principais

Segundo Jolliffe (2002), a essência da técnica de Análise de Componentes Principais é a redução da complexidade de um conjunto de dados que contém inúmeras variáveis ( $P$ ). Essa redução busca preservar ao máximo a variabilidade dos dados originais. Esse tipo de simplificação pode ser alcançada por meio da transformação dos dados em um novo conjunto de variáveis chamadas de componentes principais ( $k$ ). Tais componentes não possuem uma correlação entre si e são organizados de modo que as primeiras componentes possam reter a maior parte da variação presente no conjunto original dos dados.

O desenvolvimento matemático expresso na seção 2.2 foi extraído da obra de Johnson (2002).

Esse método tem por objetivo principal explicar a estrutura da variância/covariância dos dados através de combinações lineares das variáveis, onde as  $P$  variáveis são retidas em  $k$  componentes. Podemos escrever esses componentes como combinações lineares das  $P$  variáveis  $X_1, X_2, \dots, X_p$ . Geometricamente, essas combinações representam a seleção de um novo sistema de coordenadas obtido por meio da rotação do sistema original com

$X_1, X_2, \dots, X_p$  como o eixo de coordenadas, tais eixos representam as direções com a variabilidade máxima e nos fornecem uma descrição mais parcimoniosa da estrutura de covariância.

Seja o vetor aleatório  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  com matrix de covariância  $\Sigma$  com autovetores próprios  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p = 0$

Considere as combinações lineares:

$$Y_1 = a'_1 X = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \quad (2.2.1)$$

$$Y_2 = a'_2 X = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \quad (2.2.2)$$

$$\vdots \quad (2.2.3)$$

$$Y_p = a'_p X = a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p \quad (2.2.4)$$

que pode ser escrita como

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix}_{p \times 1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}_{p \times p} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}_{p \times 1}$$

onde

$$E(Y_i) = E(a'_i \mathbf{X}) \quad (2.2.5)$$

$$E(Y_i) = a'_i E(\mathbf{X}) \quad (2.2.6)$$

$$E(Y_i) = a'_i \mu \quad (2.2.7)$$

$$Var(Y_i) = V(a'_i \mathbf{X}) \quad (2.2.8)$$

$$Var(Y_i) = a'_i V(\mathbf{X}) a_i \quad (2.2.9)$$

$$Var(Y_i) = a'_i \Sigma a_i, \quad i = 1, 2, \dots, p \quad (2.2.10)$$

e

$$Cov(Y_i, Y_k) = V(a'_i \mathbf{X}, a'_k \mathbf{X}) \quad (2.2.11)$$

$$Cov(Y_i, Y_k) = a'_i \Sigma a_k, \quad i, k = 1, 2, \dots, p \quad (2.2.12)$$

São combinações lineares não correlacionadas de  $Y_1, Y_2, \dots, Y_p$  no qual a variância é máxima. Onde, podemos definir a partir disso:

- A 1ª componente principal será a combinação linear de  $a'_1 X$  que maximizará a  $V(a'_1 X)$  sujeita à restrição  $a'_1 a_1 = 1$
- A 2ª componente principal será a combinação linear de  $a'_2 X$  que maximizará a  $V(a'_2 X)$  sujeita à restrição  $a'_2 a_2 = 1$  e  $Cov(a'_1 X, a'_2 X) = 0$

⋮

- A  $i$ -ésimaª componente principal será a combinação linear de  $a'_i X$  que maximizará a  $V(a'_i X)$  sujeita à restrição  $a'_i a_i = 1$  e  $Cov(a'_i X, a'_k X) = 0$  para  $k < i$

Seja  $\Sigma$  a matriz de covariância associada ao vetor aleatório  $\mathbf{X}' = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p]$ , com pares de autovalores-autovetores  $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ , onde  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ , então a  $i$ -ésimaª componente principal é dada por

$$Y_i \approx e'_i X = e_{i1} X + e_{i2} X + \dots + e_{ip} X \quad i = 1, 2, \dots, p \quad (2.2.13)$$

com

$$Var(Y_i) = e'_i \Sigma e_i = \lambda_i \quad i = 1, 2, \dots, p \quad (2.2.14)$$

$$Cov(Y_i, Y_k) = e'_i \Sigma e_k = 0, \quad i \neq k \quad (2.2.15)$$

E seja,  $\mathbf{X}' = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p]$  com matriz de covariância  $\Sigma$ , e com pares de autovalores-autovetores  $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ , onde  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ , tomemos  $Y_i \approx e'_i X = e_{i1} X + e_{i2} X + \dots + e_{ip} X \quad i = 1, 2, \dots, p$ , como os componentes principais. Então

$$\sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \sum_{i=1}^p \text{Var}(X_i) = \lambda_1 + \lambda_2 + \dots + \lambda_p = \sum_{i=1}^p \text{Var}(Y_i) \quad (2.2.16)$$

Se  $Y_1 = e_1'X = Y_2 = e_2'X, \dots, Y_p = e_p'X$  são os componentes principais obtidos através da matriz de covariância  $\Sigma$ , então

$$\rho_{Y_i X_k} = \frac{e_{ik} \sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \quad i, k = 1, 2, \dots, p \quad (2.2.17)$$

são os coeficientes de correlação entre as componentes ( $Y_i$ ) e as variáveis ( $X_k$ ). E  $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$  são os pares de autovalores-autovetores para a matriz  $\Sigma$ . Alternativamente, também é possível obter a proporção total da  $k$ -ésima<sup>a</sup> componente principal por meio da relação:

$$\frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad i, k = 1, 2, \dots, p \quad (2.2.18)$$

## 2.3 Trabalhos Relacionados

O trabalho de *Billy Beane* foi um marco no que se diz respeito à análise estatística no esporte. A inserção da estatística no esporte não apenas revolucionaria a forma como as equipes de beisebol abordam suas estratégias, mas também serviu como inspiração para diversas modalidades. Nesse sentido, alguns dos trabalhos que citarei visam explorar essa relação.

Em Guimarães, Marques e Neto (2017), foi realizado um estudo com o objetivo de avaliar o desempenho de atletas de futebol das ligas Alemanha, Inglaterra, Espanha, Itália e Brasil. Consideraram 91 atributos mensurados por jogador nos últimos três anos em cada uma das ligas. O foco se deu nos atletas que atuaram como meio atacantes pela esquerda.

Ainda na área do desempenho no futebol, Abrahão et al. (2020), verificaram a aplicabilidade da Análise Fatorial por meio do cruzamento de variáveis provenientes de testes em um protocolo de detecção de talento esportivo no futebol, que incluiu medidas antropométricas, flexibilidade e força. A amostra envolveu 38 atletas da categoria sub-19 do Cuiabá Esporte Clube. Os testes realizados abrangeram o protocolo iSports para detecção de talento esportivo, avaliações antropométricas (massa corporal, estatura, Índice de Massa Corporal - IMC, Relação Cintura e Quadril - RCQ), teste de flexibilidade (Sentar e Alcançar), teste de força manual (*Hand Grip*) e teste de impulsão vertical.

Mas também há estudos que se utilizaram de simuladores como objeto de investigação, e Lima (2022) faz isso muito bem nesse estudo. Essa análise explora o papel crescente de técnicas e recursos computacionais na análise de dados no contexto do futebol, destacando a utilidade do *Football Manager*<sup>6</sup> e websites especializados como o *Transfermarkt*. O objetivo foi identificar relações entre indicadores de futebolistas que sugeriram o desenvolvimento de suas capacidades potenciais. Com uma amostra de 1309 jogadores entre 15 e 20 anos, coletados no FM Scout de 2014, o estudo utiliza técnicas de mineração de dados com o software Weka<sup>7</sup>. Os resultados apontam para a relevância de investir em jogadores com menos de 18 anos, indicando que a utilização do *Football Manager* pode ser uma ferramenta complementar valiosa para auxiliar gestores nas decisões relacionadas à identificação e contratação de jovens talentos.

No basquete, também percebe-se como a estatística é crucial para entender o desempenho de um time. Meneses, Luiz e Almeida 2016, examinaram a evolução do

<sup>6</sup><https://www.footballmanager.com/>, acessado em 07/10/2023

<sup>7</sup><https://www.cs.waikato.ac.nz/~ml/weka/>, acessado em 07/10/2023

basquete brasileiro com base nos índices de desempenho das equipes ao longo de três temporadas do Novo Basquete Brasil (de 2009 a 2012) e determinaram os valores de referência dos principais ID. Os dados foram coletados no site da Liga Nacional de Basquete. Foram analisados 726 jogos por meio de *ANOVA* de uma via. A temporada 2011-2012 se destacou por ter menos arremessos de três pontos e mais arremessos de dois pontos, uma melhoria nos lances livres e uma redução nos rebotes ofensivos. Os percentis indicam um alto nível de exigência no desempenho das equipes para os ID analisados, e foi observada uma tendência de mudança no estilo de jogo das equipes ao longo das últimas três temporadas.

## Capítulo 3 - Resultados

### 3.1 Campeonato Brasileiro de Futebol de 2022 - Série A

Popularmente conhecido como Brasileirão, neste ano foi realizada a 67<sup>a</sup> edição da principal divisão de futebol no Brasil. Foi disputada entre os dias 9 de abril e 13 de novembro, seguindo o formato de pontos corridos estabelecido em 2003, com 20 clubes disputando o campeonato desde 2006. Esta edição teve o campeão definido na 35<sup>a</sup> rodada, o Palmeiras(SP), que conquistou o seu 10<sup>o</sup> título da competição após a derrota do Internacional(RS) por 1-0 para o América Mineiro (MG).

Na parte inferior da tabela, o Juventude(RS) foi o primeiro clube rebaixado à Série B de 2023, após ser derrotado pelo Atlético Mineiro(MG) por 1-0 na 34<sup>a</sup> rodada. Avaí(SC), Ceará(CE) e Atlético Goianiense(GO) foram os outros clubes que foram rebaixados logo depois.

O Palmeiras foi o time com o melhor desempenho ofensivo, marcando 66 gols, e também teve a defesa mais sólida, sofrendo apenas 27 gols. O pior ataque e a pior defesa pertenceram ao rebaixado Juventude, com 29 e 69 gols, respectivamente. Germán Cana, jogador do Fluminense(RJ), conquistou o título de artilheiro do campeonato ao marcar 26 gols, enquanto Gustavo Scarpa, do Palmeiras, foi o líder em assistências, totalizando 13.

A Figura 2, está a classificação final do campeonato. Os clubes classificados até a 14<sup>a</sup> posição participaram das competições internacionais da América do Sul. Desde o Palmeiras até o Atlético Paranaense(PR), garantiram suas vagas na fase de grupos da Libertadores da América, enquanto o Atlético Mineiro e o Fortaleza(CE) se qualificaram para a fase pré-classificatória da Libertadores. Entre o 9<sup>o</sup> e o 14<sup>o</sup> lugares, os clubes asseguraram suas posições na Copa Sul-Americana. Notavelmente, Coritiba(PR) e Cuiabá(MT) foram os únicos clubes que não tiveram a oportunidade de disputar qualquer competição internacional no ano seguinte.

Figura 2: Tabela após a 38ª Rodada

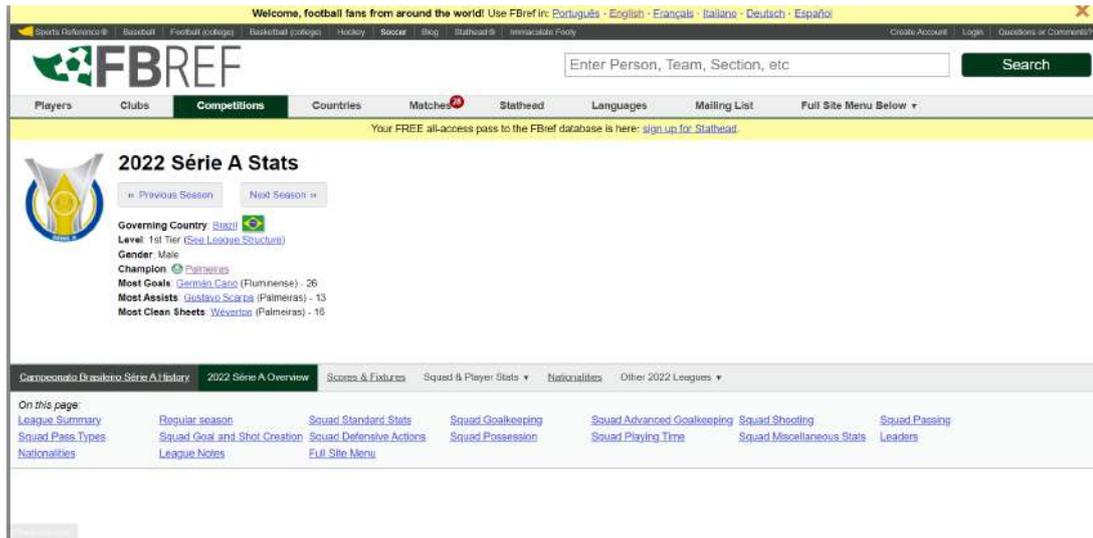
Pos	Equipe	V · D · E	Pts	J	V	E	D	GP	GC	SG
1	 Palmeiras (C)		81	38	23	12	3	66	27	+39
2	 Internacional		73	38	20	13	5	58	31	+27
3	 Fluminense		70	38	21	7	10	63	41	+22
4	 Corinthians		65	38	18	11	9	44	36	+8
5	 Flamengo		62	38	18	8	12	60	39	+21
6	 Athletico Paranaense		58	38	16	10	12	48	48	0
7	 Atlético Mineiro		58	38	15	13	10	45	37	+8
8	 Fortaleza		55	38	15	10	13	46	39	+7
9	 São Paulo		54	38	13	15	10	55	42	+13
10	 América Mineiro		53	38	15	8	15	40	40	0
11	 Botafogo		53	38	15	8	15	41	43	-2
12	 Santos		47	38	12	11	15	44	41	+3
13	 Goiás		46	38	11	13	14	40	53	-13
14	 Red Bull Bragantino		44	38	11	11	16	49	59	-10
15	 Coritiba		42	38	12	6	20	39	60	-21
16	 Cuiabá		41	38	10	11	17	31	42	-11
17	 Ceará		37	38	7	16	15	34	41	-7
18	 Atlético Goianiense		36	38	8	12	18	39	57	-18
19	 Avai		35	38	9	8	21	34	60	-26
20	 Juventude		22	38	3	13	22	29	69	-40

Fonte: Wikipédia, acessado em 07/10/2023

### 3.2 Coleta dos dados

Os dados analisados foram coletados e organizados em planilhas Excel, disponibilizados em FBref<sup>8</sup>, link para a página onde foram coletados os dados, conforme mostrado na Figura 3.

Figura 3: Pagina inicial - Campeonato Brasileiro 2022



Ao acessar essa página, role o *mouse* para baixo e você encontrará o que é mostrado na Figura 4, certifique-se de que a marcação para os dados da série A está selecionada, como na Figura 5.

Figura 4: Classificação dos clubes

Overall		Home/Away																		
Rk	Squad	MP	W	D	L	GF	GA	GD	Pts	Pts/MP	xG	xGA	xGD	xGD/90	Attendance	Top Team	Goalkeepers	Notes		
1	Palmeiras	38	23	12	3	66	27	+39	81	2.13	69.3	33.2	+36.1	+0.95	30,792			→ Copa Libertadores via league finish		
2	Internacional	38	20	13	5	58	31	+27	73	1.92	55.8	45.1	+10.8	+0.28	19,889	Alexandre Alemão - 8	Daniel	→ Copa Sudamericana via league finish		
3	Fluminense	38	21	7	10	63	41	+22	70	1.84	50.1	40.0	+9.7	+0.26	24,645	Geovane	→ Copa Sudamericana via league finish			
4	Corinthians	38	18	11	9	44	36	+8	65	1.71	40.9	41.0	-0.2	0.00	36,598	Rogério Lourenço - 10	Luísão	→ Copa Sudamericana via league finish		
5	Flamengo	38	18	8	12	60	39	+21	62	1.63	65.4	35.8	+29.6	+0.78	45,630	Pedro, Gabriel Barbosa - 11	Santos	→ Copa Libertadores via league finish		
6	Atl. Paranaense	38	16	10	12	48	48	0	58	1.53	47.1	48.4	-1.3	-0.04	16,584	De	→ Copa Libertadores via league finish			
7	Atlético Mineiro	38	15	13	10	45	37	+8	58	1.53	54.1	33.7	+20.4	+0.54	27,187	Rhulk - 12	Everson	→ Copa Libertadores via league finish		
8	Coritiba	38	15	10	13	46	39	+7	55	1.45	44.5	42.9	+1.5	+0.04	27,488	Moses, Thiago Gathardo - 7	Fernando Misuel	→ Copa Libertadores via league finish		
9	São Paulo	38	13	15	10	55	42	+13	54	1.42	49.7	44.4	+5.3	+0.14	27,157	Jonathan Calieri - 18	Jandrei	→ Copa Libertadores via league finish		
10	América (MG)	38	15	8	15	40	40	0	53	1.39	45.6	49.7	-4.1	-0.11	2,876	Hennoue Almeida - 6	Matheus Cavichin	→ Copa Sudamericana via league finish		
11	Botafogo (RJ)	38	15	8	15	41	43	-2	53	1.39	40.8	50.4	-9.6	-0.25	20,286	Erison - 7	Gatto Fernández	→ Copa Sudamericana via league finish		
12	Santos	38	12	11	15	44	41	+3	47	1.24	46.2	48.0	-1.8	-0.05	9,575	Marcos Leonardo - 13	João Paulo	→ Copa Sudamericana via league finish		
13	Ceará	38	11	13	14	40	53	-13	46	1.21	34.6	61.6	-27.0	-0.71	9,207	Pedro Baía - 19	Tadeu			
14	Bragantino	38	11	11	16	49	59	-10	44	1.16	52.7	49.5	+3.1	+0.08	4,219	Leão Cândido - 11	Cleiton Schweigber			
15	Coritiba	38	12	6	20	39	60	-21	42	1.11	42.3	53.6	-11.2	-0.30	15,495	Alef Manga - 9	Alex Huralha			
16	Cuiabá	38	10	11	17	31	42	-11	41	1.08	41.0	47.8	-6.8	-0.18	10,912	Deverson - 6	Walter			
17	Ceará	38	7	16	15	34	41	-7	37	0.97	43.9	48.0	-4.0	-0.11	27,297	Silven Mendoza - 10	João Ricardo Riedi	Relegated		
18	Atl. Goianiense	38	8	12	18	39	57	-18	36	0.95	43.1	51.2	-8.2	-0.21	5,911	Wellington Rato - 8	Ronaldo	Relegated		
19	Avaí	38	9	8	21	34	60	-26	35	0.92	44.0	62.6	-18.6	-0.49	7,776	Guilherme Bissoli - 14	Vladimir	Relegated		
20	Juventude	38	3	13	22	29	69	-40	22	0.58	33.4	57.1	-23.6	-0.62	3,200	César Ruiz, Isidro Pitta - 5	César Reis	Relegated		

<sup>8</sup><https://fbref.com/en/comps/24/2022/2022-Serie-A-Stats>, acessado em 17/09/2023

Figura 5: Marcação correta dos dados



### 3.2.1 Tipos de Variáveis

As planilhas coletadas consideram variáveis de diferentes tipos, algumas específicas para os goleiros, como *Goalkeeping* e *Advanced Goalkeeping*, enquanto as demais são específicas dos jogadores de linha: *Shooting*, com as variáveis relacionadas a chute; *Passing*, com as variáveis relativas a passe; *Pass Types*, variáveis por tipo de passe; *Goal and Shot Creation*, uma composição das variáveis que medem a criação de jogadas de ataque; *Defensive Actions*, composta pelas variáveis de defesa; *Possession*, organizada pelas variáveis que medem a quantidade por toques na bola; *Playing Time*, constituída por variáveis relativas a tempo de atuação ; e *Miscellaneous Stats*, com as demais variáveis.

Cada variável foi codificada com uma ou duas letras seguidas do número correspondente à sua posição no banco de dados no momento da coleta dos dados, sendo: Gxx(*Goalkeeping* e *Advanced Goalkeeping*), Axx(*Shooting*), Pxx(*Passing*), TPxx(*Pass Types*), GSCxx(*Goal and Shot Creation*), Dxx(*Defensive Actions*), POxx(*Possession*), Txx(*Playing Time*) e Mxx(*Miscellaneous Stats*).

### 3.2.2 Tratamento e divisão dos dados

Para lidar com os dados faltantes, optou-se por realizar a imputação utilizando a mediana. Essa escolha foi feita considerando que, para essas variáveis específicas, a distribuição dos dados apresentava assimetria, tornando a mediana uma escolha mais robusta em comparação a média.

Após a organização e padronização dos dados, procedeu-se à avaliação geral, considerando o conjunto completo dos dados, bem como às análises específicas para os primeiros e últimos colocados. Os clubes foram agrupados em dois conjuntos: o Grupo 1, contendo os times classificados para a fase de grupo da libertadores (Palmeiras, Internacional, Fluminense, Corinthians, Flamengo e Athletico Paranaense) e o Grupo 2, os clubes das quatro últimas colocações (Ceará, Atlético Goianiense, Avaí e Juventude). Foram considerados todos os jogadores que participaram da competição.

- **Defensores:** Geral com 243 jogadores, sendo 72 nos times do G6 e 52 nos times do Z4.
- **Meio Campistas:** Geral com 238 jogadores, sendo 61 nos times do G6 e 49 nos times do Z4.
- **Atacantes:** Geral com 221 jogadores, sendo 71 nos times do G6 e 46 nos times do Z4.

A divisão por área de atuação deu-se pelo fato de que, no momento da coleta, a variável "Pos", presente em todas as tabelas, classificava os jogadores por setor e não posição. Embora a alternativa de subdividir os dados com base nas posições mais tradicionais tenha sido considerada, as fontes de consulta não mostravam um consenso em relação às posições dos jogadores.

### 3.3 Defensores

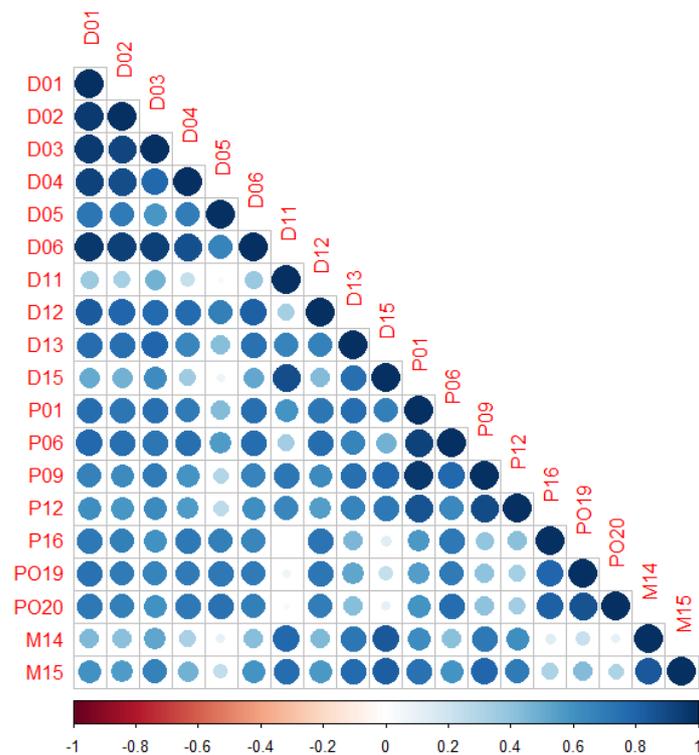
Composto pelos zagueiros e laterais, o setor defensivo desempenha um papel importante na estrutura de um time. Os zagueiros são os responsáveis pela proteção da área e pela interceptação direta dos ataques adversários. Espera-se que um zagueiro tenha boa marcação, antecipação de jogadas, jogo aéreo afiado e habilidades em saída de bola. Já os laterais ocupam uma posição mais externa, podendo contribuir tanto na defesa quanto no ataque. Além de uma boa marcação, espera-se que sejam rápidos, tenham bom passe e cruzamento, além de uma boa chegada ao ataque. Nesse sentido, foi feita uma seleção das variáveis que se adequariam a esses requisitos, conforme indicado na Tabela 1.

Tabela 1: Descrição das variáveis dos defensores - Geral

Nome	Resumo	Nome	Resumo
D01	Número de divididas	P01	Passes certos
D02	Número de posses ganhas após uma dividida	P06	Passes curtos certos
D03	Divididas no terço defensivo do campo	P09	Passes médios certos
D04	Divididas no terço central do campo	P12	Passes longos certos
D05	Divididas no terço avançado do campo	P16	Passes que geram um chute
D06	Duelos individuais ganhos	P17	Passes no terço avançado
D11	Bloqueios de chute	P19	Cruzmentos para a grande área
D12	Bloqueios de passe	PO20	Domínios no terço de ataque
D13	Número de interceptações	M14	Duelos aéreos ganhos
D15	Bolas salvas	M15	Duelos aéreos perdidos

A Figura 6 apresenta a matriz de correlação para todos os jogadores, na qual é possível avaliar como cada uma das variáveis se relaciona com as demais. A cor azul indica uma correlação positiva, enquanto tons de branco indicam uma relação inexistente. Verifica-se que D01, D02 e D03 parecem estar fortemente correlacionadas. As variáveis de ganho (M14) e perda (M15) de bolas aéreas, assim como as variáveis de passe, também parecem estar fortemente correlacionadas. Outro comportamento que segue a mesma relação das demais é entre P19 e PO20, assim como entre D11 e D15<sup>9</sup>.

Figura 6: Correlação entre as variáveis - Defensores



### 3.3.1 Aplicação de Componentes Principais

A Análise de Componentes Principais foi uma ferramenta útil para este estudo, especialmente para lidar com variáveis altamente correlacionadas. A multicolinearidade ocorre quando há alta correlação entre as variáveis, o que pode dificultar a interpretação dos resultados. A escolha pelo método de Análise de Componentes Principais nos permitiu reduzir a dimensionalidade dos dados, criando assim novas variáveis não correlacionadas, que chamamos de componentes principais. Esses componentes retêm a maior parte da

<sup>9</sup><https://soccernovo.com/what-is-a-clearance-in-soccer/>

variação original dos dados, ao mesmo tempo em que são ortogonais entre si, evitando assim a multicolinearidade.

Os autovalores mediram a quantidade de variância retida em cada componente. Observa-se que o primeiro autovalor é o maior, e os demais diminuem gradativamente. Autovalores maiores que 1 indicam que o componente principal é responsável por mais variabilidade do que as que seriam contabilizadas por uma das variáveis originais. Como os dados estão padronizados, utilizou-se um critério como ponto de corte.

Tabela 2: Análise de Componentes Principais dos defensores - Geral

<b>Dimensão</b>	<b>Autovalor</b>	<b>% de Variância Acumulada</b>
Dim.1	12.20557	64.23
Dim.2	3.450410	82.39
Dim.3	0.800312	86.61

A Tabela 2 nos mostra a quantidade de componentes, representadas por "Dimensão", os autovalores correspondentes, e a variância acumulada em cada uma delas. O comportamento esperado é que a primeira componente retenha para si a maior parte da variabilidade dos dados. Comportamento esse que foi observado na tabela. Os demais componentes, contribuem de forma gradativa. Como dito acima, busca-se componentes cujo o autovalor seja acima de 1. Neste caso, a opção pelo uso de duas componentes, que juntas explicam 82.61% da variabilidade presente nos dados é apropriada.

### 3.3.2 Autovetores

Os autovetores estão associados a cada um dos componentes e representam a direção e magnitude das mudanças nas variáveis originais, onde cada autovetor é uma combinação linear dessas variáveis e descreve a contribuição de cada uma para aquele componente. Cada combinação apresenta a contribuição das variáveis originais.

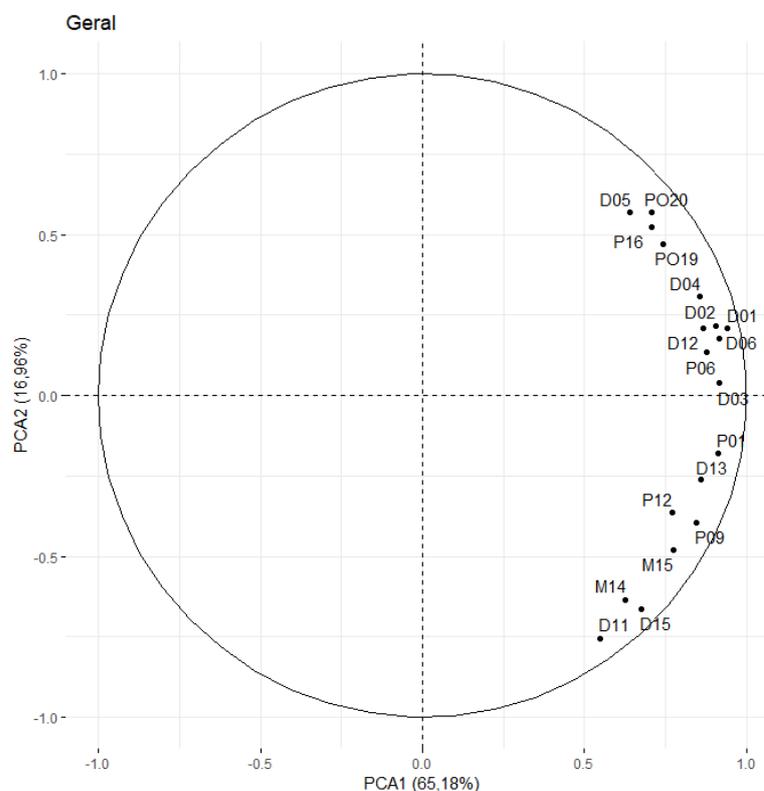
Tabela 3: Autovetores para CP1 e CP2 dos defensores

Variável	Desempenho Médio (CP1)	Fatores determinantes (CP2)
D01	0.2691	0.1122
D02	0.2590	0.1164
D03	0.2625	0.0206
D04	0.2449	0.1657
D05	0.1833	0.3057
D06	0.2620	0.0946
D11	0.1569	-0.4066
D12	0.2476	0.1127
D13	0.2455	-0.1401
D15	0.1936	-0.3573
P01	0.2605	-0.0974
P06	0.2509	0.0716
P09	0.2418	-0.2123
P12	0.2201	-0.1966
P16	0.2021	0.2810
PO19	0.2126	0.2527
PO20	0.2024	0.3064
M14	0.1791	-0.3415
M15	0.2211	-0.2580

Na Tabela 3, agora com os dados reduzidos para duas dimensões, observa-se os autovetores associados a cada variável, onde cada uma é representada nos dois componentes. Ao considerarmos um plano cartesiano, a direção da variável deve ser interpretada a partir do sinal do autovetor. Se o autovetor for positivo no componente 1 e no componente 2, a variável estará localizada no 1º quadrante. Similarmente, se ambos os autovetores forem negativos, a variável estará no 3º quadrante. Se o autovetor for positivo no componente 1 e negativo no componente 2, a variável estará no 4º quadrante; inversamente, se for negativo no componente 1 e positivo no componente 2, estará no 2º quadrante. Essa interpretação permite visualizar a orientação das variáveis nos dois primeiros componentes principais. É importante ressaltar que variáveis com sinal negativo indicam contrastes significativos, evidenciando as prioridades táticas e estratégicas adotadas no cenário atual do esporte. Essa relação pode ser vista a partir da Figura 7.

A Componente 1 pode ser entendida como uma média geral, representando características comuns que todos os jogadores compartilham em certa medida. Por outro lado, a Componente 2, devido aos contrastes observados nos coeficientes, destaca-se por diferenciar os jogadores de maneira mais específica. Variáveis como D05, P16, PO19 e PO20 destacam-se nesse sentido, indicando que os jogadores que mais participaram de divididas no terço central, realizaram passes que geraram chutes, efetuaram cruzamentos para a área adversária e dominaram no terço ofensivo se sobressaem em relação às demais. Isso mostra uma importância maior para os laterais em comparação aos zagueiros, sendo que passes médios e algumas variáveis relacionadas a divididas aparecem logo depois.

Figura 7: Variáveis dos defensores - Geral



### 3.3.3 Comparação entre os Grupos

Com o intuito de verificar se existe uma diferença significativa entre os clubes das primeiras colocações e os clubes rebaixados, inicialmente, foram analisadas ambas as matrizes de correlação.

Figura 8: Correlações dos defensores - G6

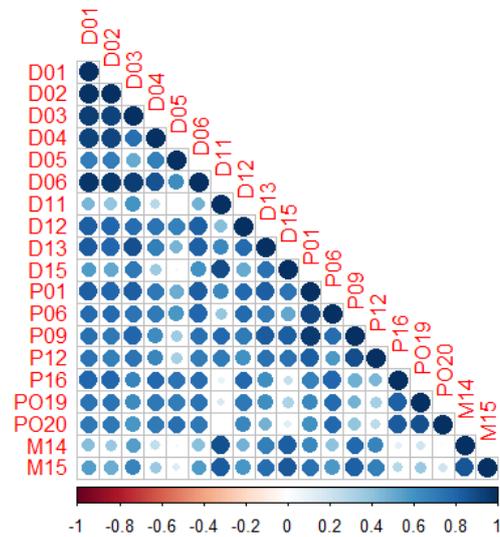
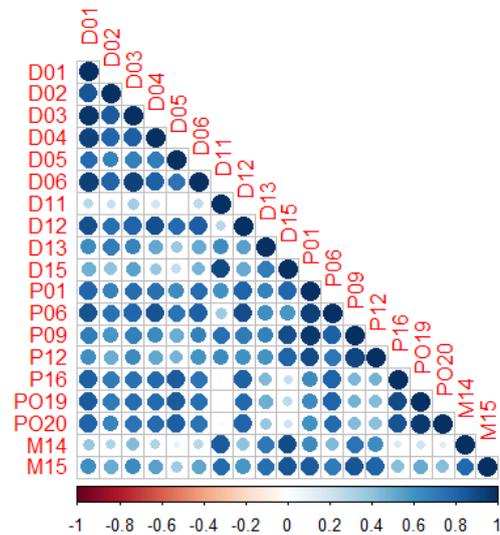


Figura 9: Correlações dos defensores - Z4



O que se percebe de imediato olhando as Figuras 8 e 9 é que a primeira possui correlações mais fortes em relação a segunda. Observa-se que, por exemplo, existe uma forte correlação entre as variáveis defensivas. As variáveis de passe também apresentam o mesmo comportamento. Isso pode significar também uma diferença na importância de cada variável em relação aos componentes.

Analogamente ao que foi visto na Tabela 2, a Tabela 4 apresenta para ambos os grupos 2 Componentes Principais, capturando 84.84% e 84.72% da variabilidade dos dados, respectivamente. Esse valor é próximo ao que foi visto para os dados gerais.

Tabela 4: Análise de Componentes Principais dos defensores - G6 e Z4

Dimensão	Autovalor G6	% Var. Acum. G6	Autovalor Z4	% Var. Acum. Z4
Dim.1	12.66265	66.64	12.53262	65.96
Dim.2	3.457765	84.84	3.565422	84.72
Dim.3	0.724469	88.65	0.700576	88.41

O que se vê na Tabela 5 é similar, se comparado à primeira componente, aos dados gerais. A segunda componente de ambos é que apresenta alguns diferenças entre os grupos. Para o Grupo 1, variáveis como D11, D15, e as relacionadas aos duelos aéreos são as que mais indicam características distintivas. Isso sugere que, nos clubes classificados para a Libertadores, jogadores com pontuações mais altas nessas variáveis destacam-se por bloqueios de chute, bolas salvas e desempenho em duelos aéreos.

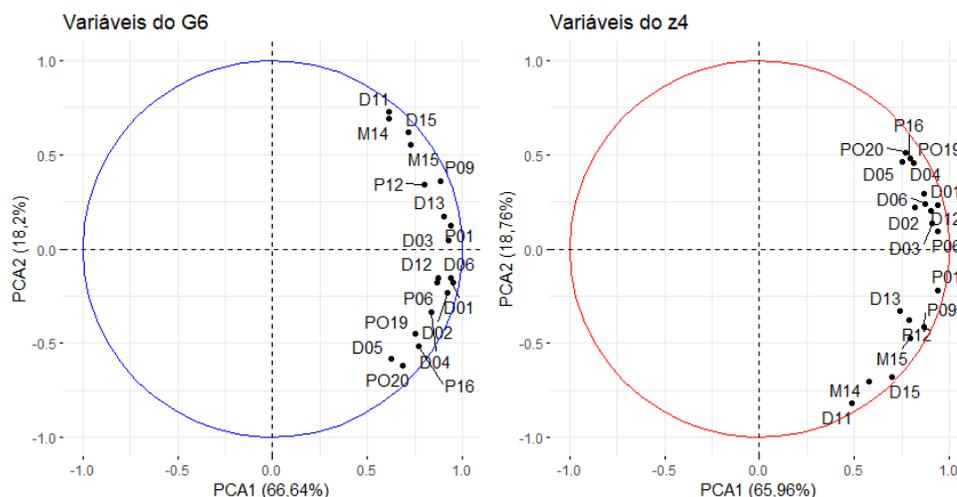
Em contraste, o Grupo 2 mostra que variáveis como P16, PO19, e PO20 são as mais influentes para a Componente 2. Isso indica que, nos clubes rebaixados, jogadores com pontuações mais altas nessas variáveis sobressaem-se em divididas no terço central, passes que geram chutes, cruzamentos para a área adversária, e domínios no terço ofensivo (O que sugere que a participação ativa dos laterais é mais determinante que a dos próprios zagueiros).

Tabela 5: Autovetores relacionados a G6 e Z4 - Defensores

Variável	Média (G6)	Média (Z4)	Determinantes (G6)	Determinantes (Z4)
D01	0.2669	0.2652	-0.0974	0.1246
D02	0.2592	0.2322	-0.1240	0.1173
D03	0.2612	0.2573	0.0250	0.0730
D04	0.2354	0.2443	-0.1790	0.1537
D05	0.1751	0.2129	-0.3130	0.2457
D06	0.2640	0.2470	-0.0827	0.1276
D11	0.1728	0.1378	0.3902	-0.4339
D12	0.2453	0.2551	-0.0812	0.1085
D13	0.2542	0.2089	0.0939	-0.1756
D15	0.2006	0.1972	0.3314	-0.3586
P01	0.2647	0.2647	0.0662	-0.1153
P06	0.2445	0.2656	-0.0972	0.0509
P09	0.2490	0.2441	0.1947	-0.2204
P12	0.2250	0.2227	0.1839	-0.2011
P16	0.2165	0.2237	-0.2790	0.2556
PO19	0.2117	0.2289	-0.2423	0.2425
PO20	0.1921	0.2185	-0.3333	0.2717
M14	0.1715	0.1626	0.3731	-0.3731
M15	0.2050	0.2248	0.2965	-0.2522

A relação da direção das variáveis, explicitada na Tabela 5 pode ser melhor visualizada na Figura 10.

Figura 10: Variáveis dos defensores - G6 e Z4



### 3.4 Meio Campistas

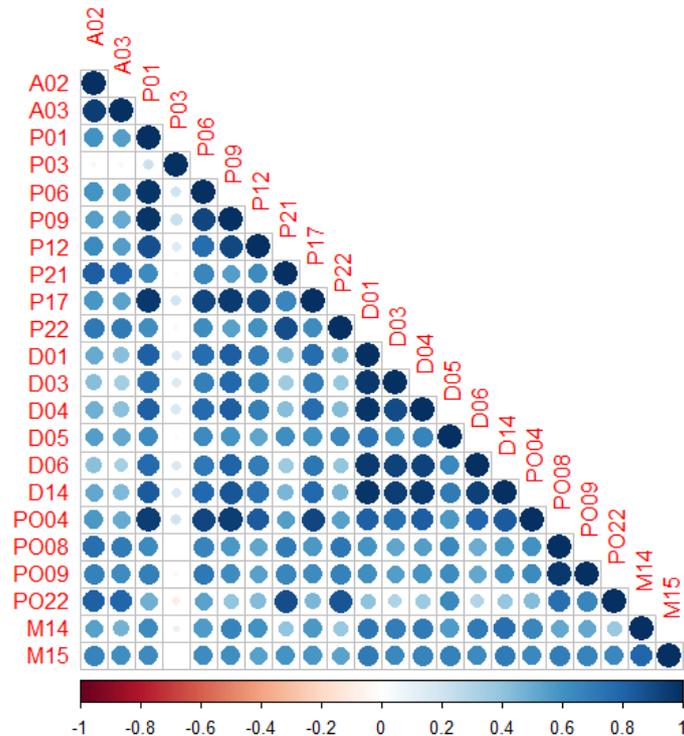
As características de um bom meio-campista variam de acordo com o esquema de cada equipe. Pode-se dividir os meio-campistas em volantes e meio-avanzados. A principal faixa de campo ocupada pelo primeiro é em frente à zaga. Esses jogadores, por característica, devem ter boa marcação e recuperação de bola, além de serem capazes de distribuir o jogo. Os volantes desempenham um papel importante na proteção da defesa e no início da construção das jogadas. E os Meio-avanzados são a parte mais criativa do meio-campo. Sua função é distribuir passes de maneira inteligente para deixar os companheiros em condições de finalizar as jogadas. Além disso, devem ter habilidades de drible para superar a marcação adversária. Espera-se também que esses jogadores possuam capacidade de finalização de média distância. Sendo assim, foram escolhidas as variáveis que mais se enquadravam nessas características.

Tabela 6: Descrição das variáveis dos meio campistas - Geral

Nome	Resumo	Nome	Resumo
A02	Finalizações	D03	Divididas no terço defensivo do campo
A03	Finalizações ao gol	D04	Divididas no terço central do campo
P01	Passes certos	D05	Duelos individuais ganhos
P03	Aproveitamento nos passes	D14	Número de intercepções
P06	Passes curtos certos	D06	Duelos individuais ganhos
P09	Passes médios certos	PO08	Tentativas de drible
P12	Passes longos certos	PO09	Tentativas de drible bem sucedidas
P21	Passes que geram um chute incluindo as assistências	P022	Passes que moveram a bola em direção do gol
P17	Passes no terço avançado do campo	PO04	Toques no terço central
P22	Cruzamentos + Passes completos para a grande área	M14	Duelos aéreos ganhos
D01	Número de divididas	M15	Duelos aéreos perdidos

A Figura 11 apresenta a correlação entre as variáveis citadas acima, notamos, onde o aproveitamento de passe não apresenta uma correlação positiva com nenhuma das demais. Tem-se uma correlação mais forte entre as variáveis de finalização, Passes certos com os passes curtos e médios, assim como com passes longos, sendo esse último um par com correlação mais fraca em relação às outras.

Figura 11: Correlação entre as variáveis - Meio Campistas



### 3.4.1 Análise Componentes Principais

Na Tabela 7 pode-se optar pelo uso de 3 componentes, porém o ganho de variabilidade a partir do segundo é de apenas 5.5%, sendo assim para manter o padrão apresentado na análise para os defensores utilizaremos 2 componentes.

Tabela 7: Análise de Componentes Principais dos meio campistas - Geral

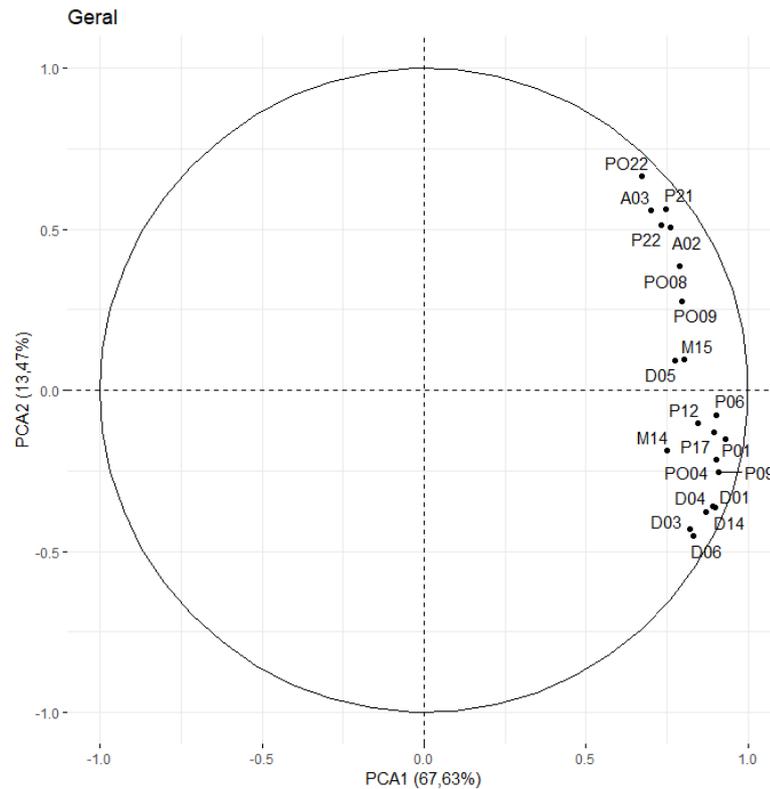
Dimensão	Autovalor	% Var. Acumulada
Dim.1	14.20389	67.63
Dim.2	2.827185	81.10
Dim.3	1.155221	86.60
Dim.4	0.683667	89.85

Assim como na Tabela 3, a primeira componente mostra uma média geral das variáveis. A componente 2, por sua vez, destaca-se por indicar as características que mais diferenciam os jogadores no contexto específico dos meio-campistas. Nota-se que variáveis como Finalizações, Finalizações ao gol, Passes que geram um chute e Passes + Cruzamentos para a grande área, além de Passes que moveram a bola em direção ao gol e Tentativas de drible evidenciam a qualidade do meio campista como fator determinante.

Tabela 8: Autovetores para CP1 e CP2 dos meio campistas

Variável	Desempenho Médio (CP1)	Fatores determinantes (CP2)
A02	0.2020	0.2999
A03	0.1852	0.3319
P01	0.2469	-0.0898
P06	0.2389	-0.0455
P09	0.2412	-0.1508
P12	0.2242	-0.0607
P21	0.1977	0.3352
P17	0.2369	-0.0769
P22	0.1939	0.3048
D01	0.2367	-0.2144
D03	0.2175	-0.2566
D04	0.2302	-0.2247
D05	0.2057	0.0547
D06	0.2202	-0.2694
D14	0.2386	-0.2170
PO04	0.2388	-0.1287
PO08	0.2095	0.2285
PO09	0.2106	0.1645
PO22	0.1778	0.3947
M14	0.1986	-0.1114
M15	0.2127	0.0569

Figura 12: Variáveis dos meio campistas - Geral



### 3.4.2 Comparação entre os Grupos

Prosseguindo com a comparação entre os grupos, as Figuras 13 e 14 apresentam as correlações específicas em cada conjunto. A primeira figura expõe padrões semelhantes aos observados no contexto geral de todos os clubes. No entanto, a segunda figura destaca uma correlação extremamente elevada nas colunas referentes ao Z4. Uma primeira suposição é que houve uma maior participação defensiva desses jogadores, e isso pode ter contribuído para esse fenômeno, considerando que a maioria das correlações distintas nas outras figuras envolvam variáveis defensivas.

Figura 13: Correlações dos meio campistas - G6

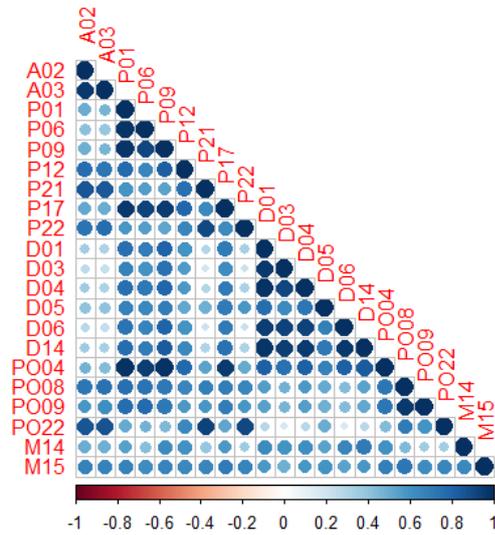
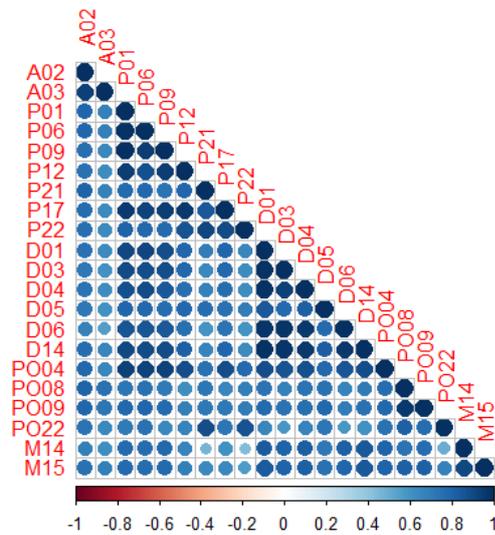


Figura 14: Correlações dos meio campistas - Z4



A Tabela 9 mostra uma relação parecida quando são avaliados o grupo 1 e o geral. Os primeiros componentes retêm para si, 64.70% e 67.63%, respectivamente. Seguidas de uma segunda componente com mais de 81% de variabilidade. Os componentes do grupo 2 apresentam um padrão diferente. A primeira explica mais de 78% da variabilidade dos dados. Um fator que pode ter contribuído para tal evento foi a alta correlação vista na Figura 14.

Tabela 9: Análise de Componentes Principais dos meio campistas - G6 e Z4

Dimensão	Autovalor G6	% Var. Acum. G6	Autovalor Z4	% Var. Acum. Z4
Dim.1	13.58794	64.70	16.51925	78.66
Dim.2	3.546074	81.59	1.503625	85.82
Dim.3	1.196285	87.28	1.065126	90.89

Na Tabela 10, analogamente ao que foi visto nas tabelas anteriores, o primeiro componente mostra a média geral das variáveis. Para o componente 2, as variáveis de finalização tem maior peso quando observamos o primeiro grupo, enquanto as demais variáveis apresentam um contribuição equivalente. Também foi realçada a qualidade do passe, quando observamos as variáveis P21, P22 e PO22. Vale ressaltar a diferença de sentido nas variáveis P17(Passes no terço avançado), D05(Duelos individuais ganhos) e M15(Duelos aéreos perdidos).

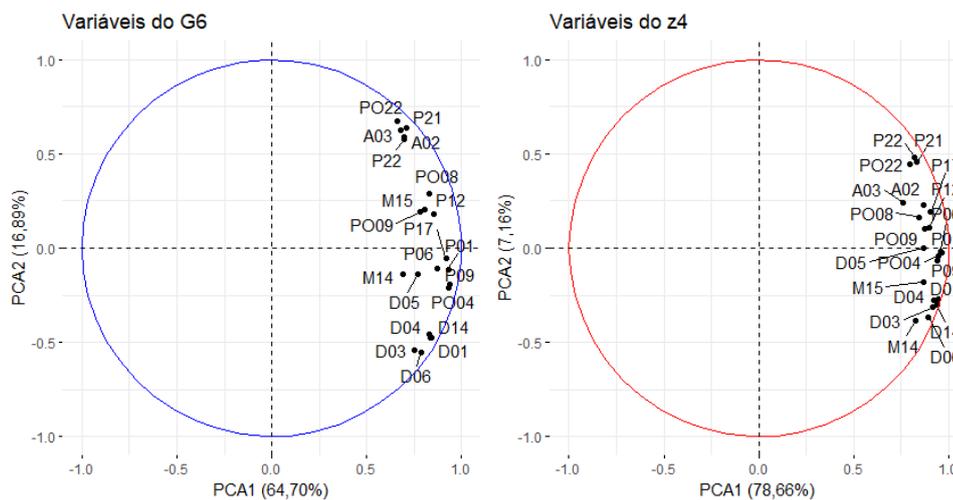
Tabela 10: Autovetores relacionados a G6 e Z4 - Meio campistas

Variável	Média (G6)	Média (Z4)	Determinantes (G6)	Determinantes (Z4)
A02	0.1883	0.2142	0.3146	0.1881
A03	0.1835	0.1872	0.3327	0.1974
P01	0.2518	0.2368	-0.0593	-0.0185
P06	0.2356	0.2357	-0.0560	-0.0134
P09	0.2535	0.2321	-0.1023	-0.0562
P12	0.2307	0.2208	0.0967	0.0904
P21	0.1924	0.2049	0.3400	0.3735
P17	0.2488	0.2226	-0.0294	0.1564
P22	0.1892	0.2013	0.3052	0.3946
D01	0.2266	0.2327	-0.2531	-0.2200
D03	0.2042	0.2254	-0.2879	-0.2570
D04	0.2248	0.2271	-0.2432	-0.2246
D05	0.2089	0.2141	-0.0746	0.0018
D06	0.2129	0.2200	-0.2950	-0.2987
D14	0.2289	0.2309	-0.2520	-0.2430
PO04	0.2532	0.2328	-0.1111	-0.0339
PO08	0.2246	0.2073	0.1539	0.1341
PO09	0.2126	0.2146	0.1006	0.0849
PO22	0.1798	0.1963	0.3591	0.3649
M14	0.1869	0.2033	-0.0729	-0.3161
M15	0.2186	0.2133	0.1083	-0.1466

Esse comportamento torna-se mais evidente ao observar a Figura 15. Observa-se que não há grandes variações quanto à direção da maioria das variáveis. Destacam-se as variáveis mencionadas anteriormente: P17, M15 e D05, que apresentam direções diferentes nos gráficos para a componente 2. Perceber-se que P17 contribuir mais para o grupo 2 em relação ao grupo 1 e ao geral, M15 ser mais notável no grupo 1 do que no grupo 2. E

D05 ser menos importante para o grupo 1 em relação aos demais.

Figura 15: Variáveis dos meio campistas - G6 e Z4



### 3.5 Atacantes

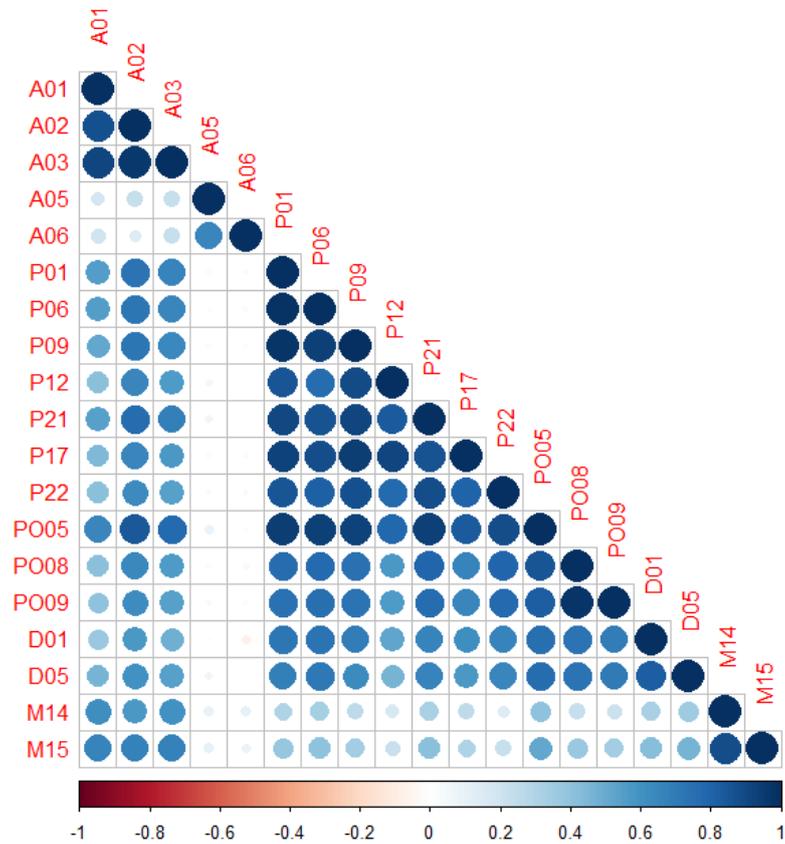
São os jogadores responsáveis por marcar gols. Atuam na linha de frente tendo como principal função finalizar as jogadas. Podem ser divididos em centroavantes, que são os jogadores que atuam mais centralizados, buscam sempre se posicionar bem para finalizar as jogadas, e os pontas, que atuam principalmente nas laterais do campo, geralmente são jogadores rápidos e habilidosos, sendo responsáveis por driblar os adversário e criar situações de gols.

Tabela 11: Descrição das variáveis dos atacantes - Geral

Nome	Resumo	Nome	Resumo
A01	Gols marcados	P17	Passes no terço avançado do campo
A02	Finalizações	P22	Cruzamentos + Passes completos para a grande área
A03	Finalizações ao alvo	P005	Toques no terço ofensivo do campo
A05	Média de finalizações por jogo	PO8	Tentativas de drible
A06	Média de finalizações certas por jogo	PO09	Tentativas de drible bem sucedidos
P01	Passes certos	D01	Número de divididas
P06	Passes curtos certos	D05	Duelos individuais ganhos
P09	Passes médios certos	M14	Duelos aéreos ganhos
P12	Passes longos certos	M15	Duelos aéreos perdido
P21	Passes que geram um chute		

Pela Figura 16, observa-se que as médias de finalizações ao gol e de finalizações por jogo não apresentam uma correlação sequer moderada com as demais variáveis. Este cenário pode ser considerado como um argumento para a consideração da remoção dessas variáveis da análise. Por outro lado, as variáveis relacionadas ao ataque parecem estar mais correlacionadas entre si, embora apresentem uma correlação moderada com as demais. No contexto dos passes, constata-se uma alta correlação entre essas variáveis e as de defesa, contrastando com uma correlação mais baixa quando comparados com o aproveitamento aéreo.

Figura 16: Correlação entre as variáveis - Atacantes



### 3.5.1 Análise dos Componentes Principais

A Tabela 12 mostra que os 2 primeiros componentes explicam mais de 80% da variabilidade dos dados, em consonância ao que foi visto para os grupos anteriores. Logo, adotou-se o uso dessas 2 componentes.

Tabela 12: Análise de Componentes Principais dos atacantes - Geral

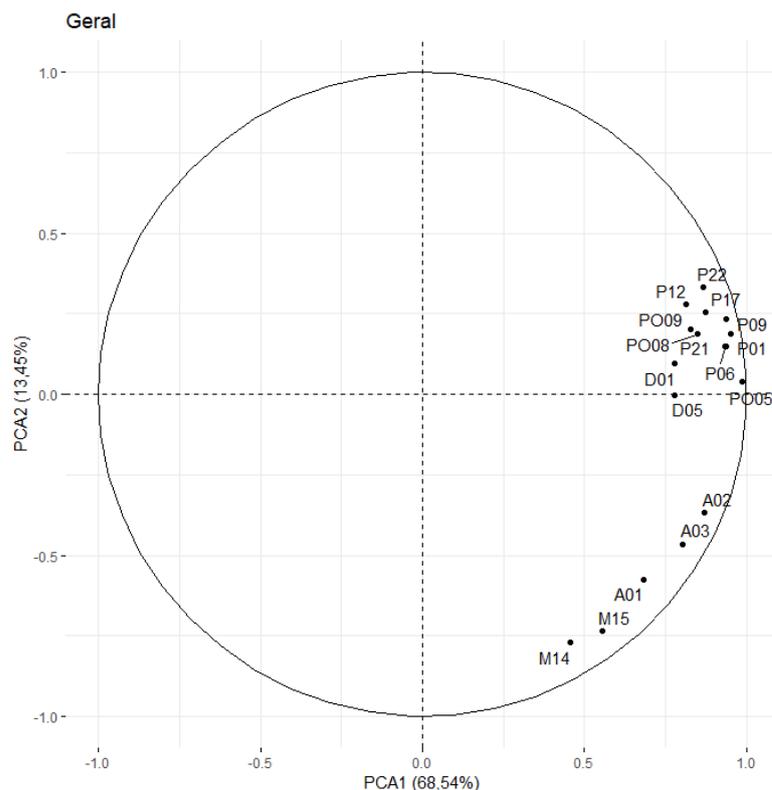
Dimensão	Autovalor	% Var. Acumulada
Dim.1	11.65245	68.54
Dim.2	2.287292	81.99
Dim.3	1.065049	88.26

A componente 1, mais uma vez, indica uma contribuição média de cada variável. Já a componente 2 destaca-se a variável P22, indicando inicialmente uma maior participação dos pontas nos cruzamentos e passes completos para a grande área, pois as variáveis A01, A02 e A03 indicam um contraste em relação a variável mencionada. Outras variáveis de passe também aparecem contribuindo significativamente, é o caso de P17 e P09. Variáveis que mostram o desempenho em duelos aéreos nos mostram que tais elementos apresentam valores negativos, indicando uma associação inversa.

Tabela 13: Autovetores para CP1 e CP2 dos atacantes

Variável	Desempenho Médio (CP1)	Fatores determinantes (CP2)
A01	0.1994	-0.3808
A02	0.2544	-0.2425
A03	0.2346	-0.3075
P01	0.2783	0.1238
P06	0.2741	0.0978
P09	0.2739	0.1545
P12	0.2377	0.1839
P21	0.2737	0.0975
P17	0.2559	0.1677
P22	0.2539	0.2190
PO05	0.2889	0.0257
PO08	0.2484	0.1240
PO09	0.2420	0.1323
D01	0.2277	0.0621
D05	0.2279	-0.0022
M14	0.1330	-0.5101
M15	0.1623	-0.4856

Figura 17: Variáveis dos atacantes - Geral



### 3.5.2 Comparação entre os grupos

O que se vê nas Figuras 18 e 19 é uma alta interação entre as variáveis de desempenho em finalizações, com destaque para a maior correlação entre gols e finalizações no primeiro grupo. Além disso, também é observável a forte relação entre participação ofensiva e passes em geral, mostrando que um bom atacante também deve construir jogadas. Em contra partida, apenas o desempenho aéreo não parece muito correlacionado com as demais

Figura 18: Correlações dos atacantes - G6

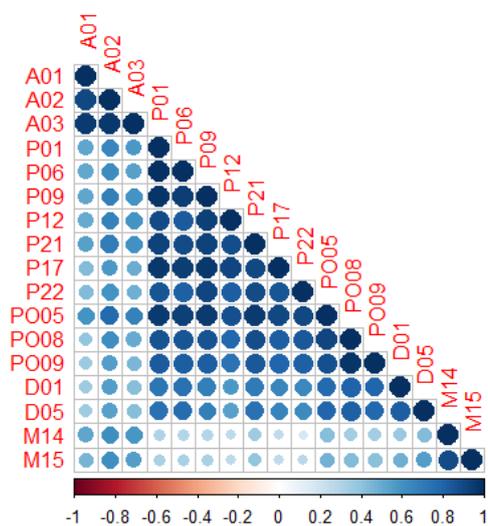
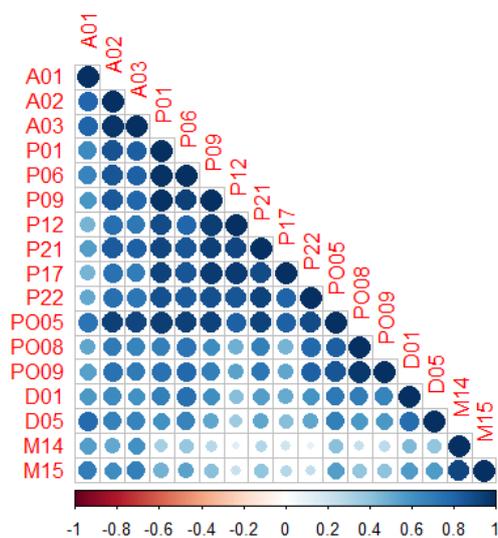


Figura 19: Correlações dos atacantes - Z4



A Tabela 14 indica um comportamento similar ao considerado em todos os jogadores. As duas primeiras componentes explicam mais de 80% da variabilidade dos dados, como visto também na Tabela 12.

Tabela 14: Análise de Componentes Principais dos atacantes - G6 e Z4

Dimensão	Autovalor G6	% Var. Acum. G6	Autovalor Z4	% Var. Acum. Z4
Dim.1	11.89517	69.97	11.93968	70.23
Dim.2	2.283106	83.40	2.177252	83.04
Dim.3	1.268185	90.86	1.141576	89.75

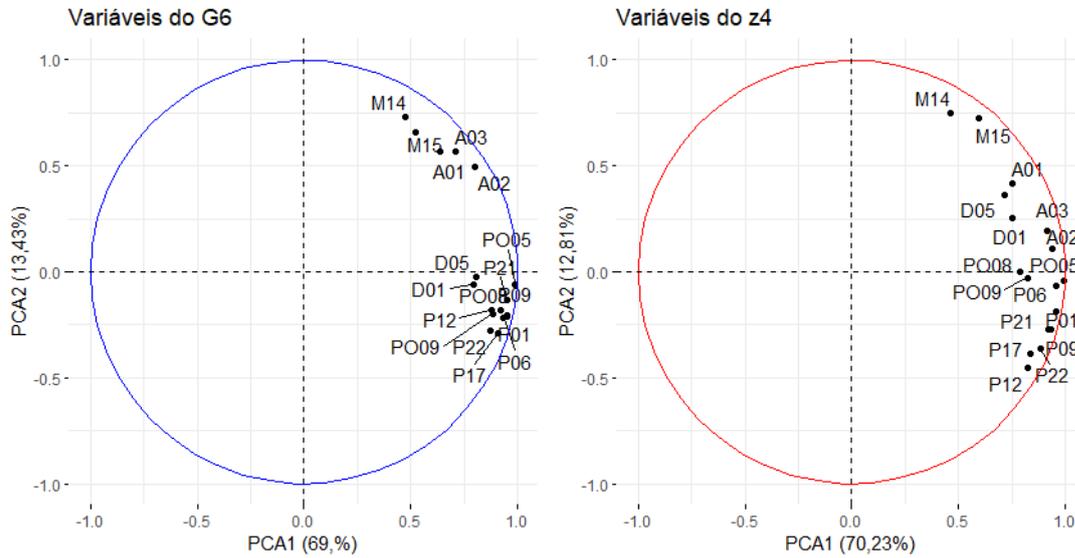
Na Tabela 15, observa-se padrões intrigantes no componente 2, os atletas do Grupo 6 demonstram uma superioridade notável em aspectos ofensivos, particularmente nas variáveis A01, A02 e A03, quando comparados aos jogadores do Grupo Z4. Contrariamente, os jogadores do Grupo Z4 exibem uma participação defensiva mais proeminente em relação ao primeiro grupo, sugerindo uma contribuição significativa nos momentos defensivos das partidas. Ambos os grupos mostram um envolvimento substancial no desempenho aéreo, indicando uma presença notável nesse aspecto do jogo.

Tabela 15: Autovetores relacionados a G6 e Z4 - Atacantes

Variável	Média (G6)	Média (Z4)	Determinantes (G6)	Determinantes (Z4)
A01	0.1854	0.2185	0.3773	0.2830
A02	0.2325	0.2719	0.3296	0.0750
A03	0.2064	0.2654	0.3753	0.1324
P01	0.2761	0.2780	-0.1407	-0.1273
P06	0.2709	0.2765	-0.1429	-0.0459
P09	0.2759	0.2698	-0.1371	-0.1861
P12	0.2555	0.2384	-0.1185	-0.3076
P21	0.2754	0.2664	-0.0891	-0.1827
P17	0.2633	0.2424	-0.1921	-0.2600
P22	0.2534	0.2555	-0.1838	-0.2450
PO05	0.2871	0.2869	-0.0385	-0.0273
PO08	0.2665	0.2283	-0.1202	-0.0002
PO09	0.2563	0.2381	-0.1306	-0.0184
D01	0.2302	0.2183	-0.0388	0.1718
D05	0.2335	0.2075	-0.0156	0.2459
M14	0.1384	0.1334	0.4822	0.5067
M15	0.1510	0.1728	0.4351	0.4903

A relação das variáveis de ataque e defesa são melhores expressas na Figura 20.

Figura 20: Variáveis dos atacantes - G6 e Z4



### 3.6 Criação dos Escores

A introdução da Análise de Componentes Principais (PCA) emerge como uma valiosa abordagem na investigação da estrutura subjacente aos dados. Os scores resultantes dessa análise proporcionam uma síntese representativa das relações entre variáveis ortogonais, viabilizando uma interpretação simplificada do desempenho dos grupos em diferentes posições. Destaco, nesta seção, os 10 primeiros colocados para cada componente, aprimorando a focalização do estudo.

Os escores, desempenham papel fundamental após a redução de  $P$  para  $k$  dimensões. Nesse contexto, os  $k$  componentes principais emergem como novos indivíduos, delineando a base para análises subsequentes. A utilização dos escores desses componentes permeia toda a análise, conferindo uma compreensão mais aprofundada da estrutura dos dados e facilitando a interpretação de relações significativas entre variáveis, refinando assim a percepção do desempenho dos grupos.

### 3.6.1 Defensores

A Tabela 5 exibiu os autovetores associados a cada nova variável, os quais desempenharão um papel crucial no cálculo da expressão correspondente a cada componente. Estes autovetores, derivados durante o processo de análise, forneceram as direções e magnitudes que orientam a contribuição de cada variável para a formação dos componentes principais. Portanto temos:

$$CP1_{G6} = 0.2669X_{D01} + 0.2592X_{D02} + \dots + 0.2050X_{M15}$$

$$CP1_{Z4} = 0.2652X_{D01} + 0.2322X_{D02} + \dots + 0.2248X_{M15}$$

$$CP2_{G6} = -0.0974X_{D01} - 0.1240X_{D02} + \dots + 0.2965X_{M15}$$

$$CP2_{Z4} = 0.1246X_{D01} + 0.1173X_{D02} + \dots - 0.2522X_{M15}$$

Tabela 16: TOP 10 Melhores defensores (CP1) - G6

Nome	Time	Score
Marcos Rocha	Palmeiras	8.8736
Renê	Internacional	8.6451
Ayrton Lucas	Flamengo	6.7844
Murilo	Palmeiras	5.9682
Joaquín Piquerez	Palmeiras	5.9470
Gustavo Gómez	Palmeiras	5.5760
Matheuzinho	Flamengo	5.4627
Fabricao Bustos	Internacional	5.3991
Samuel Xavier	Fluminense	5.2889
Pedro Henrique	Athlético Paranaense	4.8267

Tabela 17: TOP 10 Melhores defensores (CP1) - Z4

<b>Nome</b>	<b>Time</b>	<b>Score</b>
Kevin Avaí	Avaí	9.5042
Bruno Cortez	Avaí	7.0032
Nino Paraíba	Ceará	6.9274
Rodrigo Alves Soares	Juventude	6.8536
Jefferson	Atlético Goianiense	6.2297
Luis Eduardo	Atlético Goianiense	4.4307
Paulo Henrique	Juventude	3.8991
Luiz Otávio	Ceará	3.7370
Wanderson	Atlético Goianiense	3.0199
Rafael Forster	Juventude	2.6746

Tabela 18: TOP 10 Melhores defensores (CP2) - G6

<b>Nome</b>	<b>Time</b>	<b>Score</b>
Gustavo Gómez	Palmeiras	5.2002
Vitão	Internacional	4.9892
Murilo	Palmeiras	4.1358
Pedro Henrique	Athlético Paranaense	3.6069
Gil	Corinthians	2.9372
Manoel	Fluminense	2.8321
Pablo	Flamengo	2.7796
Nino	Fluminense	2.6996
Gabriel Mercado	Internacional	2.5342
Rodrigo Moledo	Internacional	2.0810

Tabela 19: TOP 10 Melhores defensores (CP2) - Z4

Nome	Time	Score
Paulo Henrique	Juventude	3.7366
Nino Paraíba	Ceará	2.9969
Natanael	Avaí	2.4374
Bruno Cortez	Avaí	2.3742
Arthur Henrique	Atlético Goianiense	2.3547
Kevin	Avaí	1.8630
Luis Eduardo	Atlético Goianiense	1.8135
Hayner	Atlético Goianiense	1.4352
Rodrigo Alves Soares	Juventude	1.4348
Renato	Avaí	1.2976

### 3.6.2 Meio Campistas

Da Tabela 10 obtem-se os seguintes componentes.

$$CP1_{G6} = 0.1883X_{A02} + 0.1835X_{A03} + \dots + 0.2186X_{M15}$$

$$CP1_{Z4} = 0.2142X_{A02} + 0.1872X_{A03} + \dots + 0.2133X_{M15}$$

$$CP2_{G6} = 0.3146X_{A02} + 0.3327X_{A03} + \dots + 0.1083X_{M15}$$

$$CP2_{Z4} = 0.1881X_{A02} + 0.1974X_{A03} + \dots - 0.1466X_{M15}$$

Tabela 20: TOP 10 Melhores meio campistas (CP1) - G6

Jogador	Clube	Score
André	Fluminense	8.7023
Carlos de Pena	Internacional	7.0393
Gustavo Scarpa	Palmeiras	6.7231
Ganso	Fluminense	6.5023
Danilo	Palmeiras	6.3476
Zé Rafael	Palmeiras	6.2297
Gabriel	Internacional	5.6923
Du Queiroz	Corinthians	5.0596
João Gomes	Flamengo	5.0558
Everton Ribeiro	Flamengo	5.0373

Tabela 21: TOP 10 Melhores meio campistas (CP1) - Z4

Jogador	Clube	Score
Raniele	Avaí	9.3338
Marlon Freitas	Atlético Goianiense	8.6737
Bruno Silva	Avaí	8.2286
Richard	Ceará	8.0136
Jadson	Juventude	7.7132
Chico	Juventude	7.0270
Gabriel Baralhas	Atlético Goianiense	6.5382
Jorginho	Atlético Goianiense	6.2022
Eduardo de Biasi	Avaí	5.1835
Richardson	Ceará	4.9426

Tabela 22: TOP 10 Melhores meio campistas (CP2) - G6

Jogador	Clube	Score
Gustavo Scarpa	Palmeiras	8.4592
David Terans	Atlético Paranaense	3.9961
Alan Patrick	Internacional	3.3883
Giorgian De Arrascaeta	Flamengo	3.2656
Everton Ribeiro	Flamengo	1.9743
Raphael Veiga	Palmeiras	1.6362
Vitor Bueno	Atlético Paranaense	1.3763
Nathan	Fluminense	1.3017
Taison	Internacional	1.1268
Eduard Atuesta	Palmeiras	1.0562

Tabela 23: TOP 10 Melhores meio campistas (CP2) - Z4

Jogador	Clube	Score
Jorginho	Atlético Goianiense	3.3626
Marlon Freitas	Atlético Goianiense	2.7215
Chico	Juventude	2.0471
Felipe Pires	Juventude	1.5735
Yuri Oliveira Lima	Juventude	1.4864
Kelvin	Atlético Goianiense	1.4594
Eduardo de Biasi	Avaí	1.4288
Jadson	Juventude	1.2526
Bruno Nazário	Juventude	1.1635
Jean Pyerre	Avaí	1.0469

### 3.6.3 Atacantes

E por fim, da Tabela 15 obtem-se os seguintes componentes.

$$CP1_{G6} = 0.1854X_{A01} + 0.2325X_{A02} + \dots + 0.1510X_{M15}$$

$$CP1_{Z4} = 0.2185X_{A01} + 0.2719X_{A02} + \dots + 0.1728X_{M15}$$

$$CP2_{G6} = 0.3773X_{A01} + 0.3296X_{A02} + \dots + 0.4351X_{M15}$$

$$CP2_{Z4} = 0.2830X_{A01} + 0.0750X_{A02} + \dots + 0.4903X_{M15}$$

Tabela 24: TOP 10 Melhores atacantes (CP1) - G6

Jogador	Clube	Score
Jhon Arias	Fluminense	13.6314
Dudu	Palmeiras	10.1673
Roger Guedes	Corinthians	7.4178
Germán Cano	Fluminense	4.9431
Gabriel Barbosa	Flamengo	4.8527
Alexandre Alemão	Internacional	4.7802
Tomás Cuello	Athletico Paranaense	4.6934
Matheus Martins	Fluminense	4.6211
Everton Soares	Flamengo	4.4570
Adson	Corinthians	3.7517

Tabela 25: TOP 10 Melhores atacantes (CP1) - Z4

Jogador	Clube	Score
Wellington Rato	Atlético Goianiense	9.4648
Vinícius	Ceará	8.4954
Stiven Mendoza	Ceará	6.8942
William Pottker	Avaí	5.5329
Lima	Ceará	5.5087
Diego Churín	Atlético Goianiense	4.7605
Guilherme Bissoli	Avaí	4.5207
Shaylon	Atlético Goianiense	3.4554
Airton	Atlético Goianiense	3.2583
Isidro Pitta	Juventude	3.2103

Tabela 26: TOP 10 Melhores atacantes (CP2) - G6

Jogador	Clube	Score
Germán Cano	Fluminense	5.7478
Rony	Palmeiras	4.4828
Alexandre Alemão	Internacional	3.4286
Pablo	Athlético Paranaense	3.2534
Yuri Alberto	Corinthians	3.2015
Pedro	Flamengo	2.8252
Vitor Roque	Athlético Paranaense	2.6024
Rômulo	Athlético Paranaense	1.4776
Rafael Navarro	Palmeiras	1.4507
Pedro Henrique	Internacional	1.2449

Tabela 27: TOP 10 Melhores atacantes (CP2) - Z4

<b>Jogador</b>	<b>Clube</b>	<b>Score</b>
Diego Churín	Atlético Goianiense	5.1004
Guilherme Bissoli	Avaí	3.6243
Cléber	Ceará	2.6824
Isidro Pitta	Juventude	2.0475
Vitor Gabriel	Juventude	1.8893
Ricardo Bueno	Juventude	1.3170
Zé Roberto	Ceará	0.9361
Stiven Mendoza	Ceará	0.8834
Matheus Peixoto	Ceará	0.7293
Airton	Atlético Goianiense	0.5649

## Conclusões

A análise aplicada aos dados do Campeonato Brasileiro de 2022 foi empregada para investigar a estrutura subjacente no desempenho dos jogadores. Os dados foram segmentados em três grupos distintos, posteriormente subdivididos de acordo com a classificação. A escolha de 2 componentes revelou-se eficaz na avaliação dos autovalores e na quantificação da variabilidade retida por eles. Além disso, a geração dos escores proporcionou uma visão mais clara do desempenho real de cada um dos 10 melhores jogadores em cada componente.

A análise dos defensores proporcionou insights valiosos sobre a relação entre as variáveis de presença ofensiva, sugerindo uma participação mais ativa dos laterais. Esse achado destaca a importância de um bom lateral não apenas na defesa, mas também na capacidade de distribuir e construir jogadas, evidenciando a versatilidade necessária para esses jogadores. Além disso, ao examinar o Componente 2, observou-se que os jogadores do G6 demonstraram um desempenho defensivo ainda melhor do que os do Z4. Essa diferença sugere que os zagueiros dos clubes classificados diretamente para a Libertadores desempenharam um papel mais proeminente do que o segundo grupo. Essa constatação reforça a crucial importância de uma defesa sólida não apenas para proteger o gol, mas também para construir as bases do sucesso de uma equipe, indicando que uma boa defesa não apenas se limita a evitar gols, mas também desempenha um papel fundamental na construção do caminho para a vitória.

Os meio-campistas revelaram algumas discrepâncias ao compararmos as direções de algumas variáveis nos dois grupos, especialmente no contexto da segunda componente. Contudo, um ponto comum identificado em ambos foi a ênfase na qualidade de passe e a valorização da capacidade de finalização. Essa observação reflete a ideia fundamental de que um meio-campista de qualidade deve primordialmente focar na criação de jogadas, enquanto, sempre que possível, demonstrar habilidades eficazes de finalização

Quanto aos atacantes, observamos algumas disparidades, especialmente no componente 2, que destacou uma qualidade ofensiva superior por parte dos jogadores do Grupo 1. Isso sugere que esses atacantes foram mais eficientes nesse aspecto específico. Por outro lado, os jogadores do Grupo 2, no componente 2, evidenciaram uma participação mais proeminente em situações defensivas. Essa observação sugere, por exemplo, uma superioridade qualitativa dos atacantes do G6 em comparação com o segundo grupo.

Por fim, ao empregar os escores, pode-se ver como os jogadores pontuaram em

cada componente, simplificando assim a avaliação de seu desempenho global. Essa abordagem proporciona uma visão mais clara e objetiva das contribuições individuais dos jogadores para os diferentes aspectos do jogo.

## Referências

- ABRAHAO, A. A. d. O. et al. Análise fatorial para detecção do talento esportivo em jogadores de futebol. Universidade Federal de Mato Grosso, 2020.
- ANDERSON, C.; SALLY, D. *The numbers game: Why everything you know about soccer is wrong*. [S.l.]: Penguin, 2013.
- GUIMARAES, C. B.; MARQUES, J. M.; NETO, A. C. Análise estatística para avaliar o desempenho de atletas de futebol profissional. Simpósio de Métodos Numéricos em Engenharia, 2017.
- JOHNSON, R. A.; WICHERN, D. W. et al. *Applied multivariate statistical analysis*. Prentice hall Upper Saddle River, NJ, 2002.
- JOLLIFFE, I. T. *Principal component analysis for special types of data*. [S.l.]: Springer, 2002.
- LEWIS, M. *Moneyball: The art of winning an unfair game*. [S.l.]: WW Norton & Company, 2004.
- LIMA, E. M. R. O uso do jogo eletrônico football manager e do portal transfermarket para mensurar o potencial de futebolistas promissores. Universidade Estadual Paulista, 2022.
- MENESES, L. R.; JUNIOR, L. E. M. G.; ALMEIDA, M. B. Análise do desempenho do basquetebol brasileiro ao longo de três temporadas do novo basquete brasil. *Revista Brasileira de Ciências do Esporte*, 2016.