



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO CIÊNCIA EXATA E DA TERRA
DEPARTAMENTO DE MATEMÁTICA
MESTRADO PROFISSIONAL EM MATEMÁTICA EM
REDE NACIONAL – PROFMAT



JOÃO VICTOR RODRIGUES DA COSTA

A UTILIZAÇÃO DO ESPORTE NO ENSINO DA ESTATÍSTICA: O CASO PRÁTICO DO FUTEBOL

Orientador:

Prof. Dr. Edgar Silva Pereira

Natal/RN - 2021

JOÃO VICTOR RODRIGUES DA COSTA

**A UTILIZAÇÃO DO ESPORTE NO
ENSINO DA ESTATÍSTICA: O CASO
PRÁTICO DO FUTEBOL**

Dissertação apresentada ao Corpo Docente do Mestrado Profissional em Matemática em Rede Nacional - PROFMAT - CCET - UFRN, como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador:

Prof. Dr. Edgar Silva Pereira

Natal/RN - 2021

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
DEPARTAMENTO DE MATEMÁTICA
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL - PROFMAT**

JOÃO VICTOR RODRIGUES DA COSTA

A UTILIZAÇÃO DO ESPORTE NO ENSINO DA ESTATÍSTICA: O CASO PRÁTICO DO FUTEBOL

Comissão Examinadora:

Prof^o. Dr^o. Edgar Silva Pereira (UFRN - Orientador)

Prof^o. Dr^o. Edgar Silva Pereira (UFRN - Membro interno)

Prof^o. Dr^o. Jaques Silveira Lopes (UFRN - Membro interno)

Prof^o. Dr^o. Wenia Valdevino Félix de Lima (UEPB - Membro externo)

Natal/RN - 2021

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Setorial Prof. Ronaldo Xavier de Arruda - CCET

Costa, João Victor Rodrigues da.

A utilização do esporte no ensino da Estatística: o caso prático do futebol / João Victor Rodrigues da Costa. - 2021. 93f.: il.

Dissertação (mestrado) - Universidade Federal do Rio Grande do Norte, Centro de Ciências Exatas e da Terra, Departamento de Matemática, Mestrado Profissional em Matemática em Rede Nacional - PROFMAT. Natal, 2021.

Orientador: Prof. Dr. Edgar Silva Pereira.

1. Estatística - Dissertação. 2. Futebol - Dissertação. 3. Comparação - Dissertação. 4. Excel - Dissertação. I. Pereira, Edgar Silva. II. Título.

RN/UF/CCET

CDU 311

Dissertação de Mestrado sob o título “A utilização do esporte no ensino da estatística: o caso prático do futebol” apresentado por João Victor Rodrigues da Costa e aceito pelo Programa de Pós-Graduação em Matemática em Rede Nacional – PROFMAT da Universidade Federal do Rio Grande do Norte, como requisito parcial para obtenção do grau de Mestre, aprovado por todos os membros da banca examinadora abaixo especificada:

Prof. Dr. Edgar Silva Pereira

Orientador

UFRN - Universidade Federal do Rio Grande do Norte

Prof. Dr. Jaques Silveira Lopes

Examinador Interno

UFRN - Universidade Federal do Rio Grande do Norte

Prof. Dr. Wenia Valdevino Félix de Lima

Examinador Externo

UEPB - Universidade Estadual da Paraíba

Agradecimentos

Quero agradecer primeiro a Deus, por me dar forças e determinação para buscar meus sonhos, profissionais e acadêmicos, além de uma melhor formação, sempre com o objetivo de aprimoramento profissional, além de toda força e saúde, para que eu tenha chegado até aqui.

A posteriori quero agradecer à minha família, em especial meus pais e minhas irmãs, por toda criação e discernimento que possibilitaram eu ser o que sou hoje.

Também quero agradecer a minha namorada, Cecília, pela parceria e companheirismo todos os dias.

Quero agradecer meus amigos de infância e do IFRN, por terem feito parte de todo meu crescimento acadêmico, cada um deles sabe da sua importância em meu processo, esta conquista é nossa.

Quero agradecer aos meus amigos do mestrado, sem os quais eu não teria chegado até aqui, esta conquista não é só minha.

Por fim, quero agradecer toda equipe de professores que puderam fazer com que esse sonho se realizasse, com o conhecimento como também no apoio moral, desde a primeira disciplina, até a defesa da dissertação. Destaque para meu professor orientador, Edgar.

Dedicatória

Dedico este trabalho a minha família, amigos, namorada e todos os profissionais que contribuíram com minha trajetória acadêmica, desde a escolha do curso até a apresentação da dissertação.

Resumo

Este trabalho tem como objetivo de apresentar alguns conceitos de Estatística descritiva e utilizá-la como uma ferramenta de comparação no contexto do futebol, com o uso do editor de planilhas Microsoft Excel. O estudo também servirá para satisfazer a curiosidade de quem gosta do esporte, pois mostrará comparações entre Campeonatos Brasileiros de Futebol da série A aos longos dos anos, por meio de gráficos e das medidas de posição e de variação. Esses resultados poderão ser vistos como previsões e estimativas médias do que ocorre nos campeonatos, o que contribui para traçar as metas e os objetivos dos times. A dissertação também apresentará como a estatística estará envolvida nesse contexto, ao mostrar padrões nas ocorrências dos campeonatos. Para embasamento teórico foram usadas definições, tabelas, gráficos, medidas de tendência central, de dispersão e analisou-se como eles podem ser utilizados. Outro objetivo desse estudo é fazer com que os alunos, guiados pelo docente, trabalhem a ideia de coleta dos dados, análise e apresentação, a fim de fazer comparações estatísticas dentro de uma situação parecida, como por exemplo em um outro esporte. Como ferramenta para o desenvolvimento do trabalho tem-se o Microsoft Excel, que além de ser um software conhecido por muitas pessoas, é de fácil uso e acesso, o que auxilia na criação de tabelas, gráficos e cálculo das medidas.

PALAVRAS-CHAVE: Estatística, Futebol, Comparação, Excel.

Abstract

This work aims to present some concepts of descriptive statistics and use it as a comparison tool, using Microsoft Excel, in the context of football. The work will also serve to satisfy the curiosity of those who like the sport, as it will show comparisons between Brazilian Soccer Championships of the A series over the years, through graphics and measurements of position and variation. These results can be seen as average predictions and estimates of what happens in the championships, and can be used to set goals and objectives for the teams. The dissertation will also show how statistics is involved in this context, showing patterns in the occurrences in championships, with a theoretical basis, using definitions, tables, graphs, measures of central tendency, dispersion and how they can be used. One of the goals is also to make students, guided by the Professor, work on the idea of data collection, analysis and presentation, seeking to make statistical comparisons within a similar situation, such as in another sport. As a tool for the development of the work we will have Microsoft Excel, a software known by many people, easy to use and access, which helps in the creation of tables, graphs and calculation of measures.

KEY-WORD: Statistics, Soccer, Comparison, Excel.

Lista de Figuras

2.1	Classificação das variáveis.	21
2.2	Gráfico da pontuação dos campeões brasileiros de 2015 até 2019 . .	25
2.3	Gráfico do resultado do Flamengo no campeonato brasileiro de 2019	26
2.4	Gráfico do desempenho dos times do campeonato brasileiro de 2019, destaque para os 4 primeiros colocados.	27
2.5	Pontuação acumulada do Vasco nas 8 primeiras rodadas	27
2.6	Gráfico em radar de Cristiano Ronaldo no FIFA 18	28
2.7	Número de vitória dos times da série A de 2019.	36
3.1	Cálculo das medidas por meio do excel	44
3.2	Comparação entre o 16º e o 17º colocado	48
3.3	Posição dos campeões brasileiros de 2016-2020	50
3.4	Pontuação acumulada do Flamengo de 2019 e Avaí de 2019	51
3.5	Pontuação acumulada do Flamengo de 2019 e Santos de 2019 . . .	52
3.6	Rendimento dos clubes dentro e fora de casa	60
4.1	Aplicativo do Excel no computador	64
4.2	Célula no Excel	65
4.3	Rendimento dos últimos 5 campeões jogando em casa	66
4.4	Inserir gráfico	67
4.5	Gráficos recomendados	68
4.6	Aproveitamento dos clubes jogando em casa	69
4.7	Aproveitamento do Flamengo, no ano de 2020, jogando em casa . .	70
4.8	Comparação do número de vitória dos últimos 5 campeões	70
4.9	Cálculo da média	71

4.10	Cálculo da mediana	72
4.11	Cálculo do desvio padrão	73
4.12	Cálculo da variância	73
1	Tabela do campeonato brasileiro de 2006	79
2	Tabela do campeonato brasileiro de 2007	80
3	Tabela do campeonato brasileiro de 2008	81
4	Tabela do campeonato brasileiro de 2009	82
5	Tabela do campeonato brasileiro de 2010	83
6	Tabela do campeonato brasileiro de 2011	84
7	Tabela do campeonato brasileiro de 2012	85
8	Tabela do campeonato brasileiro de 2013	86
9	Tabela do campeonato brasileiro de 2014	87
10	Tabela do campeonato brasileiro de 2015	88
11	Tabela do campeonato brasileiro de 2016	89
12	Tabela do campeonato brasileiro de 2017	90
13	Tabela do campeonato brasileiro de 2018	91
14	Tabela do campeonato brasileiro de 2019	92
15	Tabela do campeonato brasileiro de 2020	93

Lista de Tabelas

2.1	Classificação do campeonato brasileiro de futebol de 2019.	22
2.2	Gols marcados pelos times da série A de 2019.	24
2.3	Gols por jogo do Palmeiras na série A 2020	34
2.4	Altura dos jogadores do Atlético MG.	35
3.1	Pontuação dos campeões brasileiros	43
3.2	Pontuação dos 17º colocados	46
3.3	Pontuação dos 16º colocados	47
3.4	Pontuação dos 4º colocados	49
3.5	Número de gols marcados	53
3.6	Número de gols sofridos pelos campeões	54
3.7	Saldo de gols dos campeões	55
3.8	Jogos sem sofrer gols dos campeões (Clean sheets)	56
3.9	Aproveitamento dos campeões jogando em casa	57
3.10	Aproveitamento dos campeões jogando fora de casa	58
3.11	Aproveitamento dos campeões	59

Sumário

1	Introdução	15
2	Conceitos Fundamentais da Estatística	19
2.1	Noções básicas	19
2.1.1	População	19
2.1.2	Amostra	20
2.1.3	Censo	20
2.1.4	Parâmetro	20
2.1.5	Variável	20
2.1.6	Tabela	22
2.2	Distribuição de frequência e intervalo de classe	23
2.3	Gráficos	25
2.3.1	Gráfico de barras ou colunas	25
2.3.2	Gráfico de setores	26
2.3.3	Gráficos de segmentos ou linha	26
2.3.4	Gráfico em área	27
2.3.5	Gráfico em radar	28
2.4	Medidas de posição	29
2.4.1	Média Aritmética	29
2.4.2	Média Ponderada	30
2.4.3	Média Geométrica	30
2.4.4	Média Harmônica	31
2.4.5	Mediana	32
2.4.6	Moda	35

2.5	Medidas de dispersão	37
2.5.1	Amplitude total	38
2.5.2	Desvio Médio	38
2.5.3	Variância	39
2.5.4	Desvio Padrão	39
3	Comparação Estatística	41
3.1	Série A do Campeonato Brasileiro de Futebol	41
3.2	Comparações Estatísticas	42
3.2.1	Campeões Brasileiros	43
3.2.2	Clubes rebaixados para a Série B	45
3.2.3	O grupo dos G4	48
3.2.4	Comparação entre campeões	50
3.2.5	Número de gols marcados	53
3.2.6	Aproveitamento dos clubes	56
3.3	Sugestão de atividade	61
4	Utilização do Excel	63
4.1	O Excel	63
4.2	Tabelas	66
4.3	Gráficos	67
4.4	Medidas de tendência central	71
4.5	Medidas de dispersão	72
5	Conclusão	74
	Apêndice	79

Capítulo 1

Introdução

Hodiernamente, é comum ter dados registrados em vários aplicativos, sites e blogs, pois a tecnologia está em quase tudo, e faz com que esses dados sirvam de base para uma tomada de decisão, que pode gerar a diminuição de gastos ou, em uma empresa, o corte de funcionários. Isso pode ser decidido baseando-se nos dados que mostram o que acontece e o que poderá acontecer, uma aplicação que é análoga no esporte.

Antes de tudo é preciso entender o que é a estatística, que, segundo Correa ([4], p.8), é a ciência que fornece métodos para coleta, organização, descrição, análise e interpretação de dados, com objetivo de utilizá-los para tomadas de decisão. Wheelan ([26], p.7) diz que a Estatística nos ajuda a entender tudo, desde teste de DNA até loterias, inserida no nosso cotidiano, fica mais fácil ter um entendimento da sua importância. Existe 3 grandes áreas da estatística, a descritiva, a inferencial e a probabilística, a que vamos trabalhar é a *descritiva*. Nela são sintetizados os valores de um conjunto de números, que fazem uma representação por meio de tabelas, gráficos ou medidas descritivas. A inferencial é o processo de obter informações da amostra e a probabilística traz um modelo que estuda fenômenos da Estatística por meio de experimentos. Entretanto, Guedes [12] também enfatiza o perigo no equívoco do uso das tecnologias em comparação estatística, como o cálculo, que, sem o devido conhecimento, ou aprofundamento teórico, pode causar uma conclusão equivocada. Torna-se também importante destacar que o autor apresenta o importante uso da tecnologia na estatística, desde que seja utilizada da maneira correta.

Muitos esportes, em diversos países, usam a Estatística como ferramenta e, em alguns casos, possuem um setor específico responsável. Nesse contexto, podemos ter

como exemplo o time Liverpool, da Inglaterra, treinado por Jurgen Klopp juntamente a Ian Graham, doutor em Física pela Universidade de Cambridge. Segundo o jornal estadunidense, The New York Times, o time conquistou a Champions League com o uso da tecnologia e dados estatísticos. Ao perceber que não poderia enfrentar financeiramente - em contratações - os grandes times da Europa, usou a matemática ao seu favor: analisou mais de 100 mil dados estatísticos para a contratação de jogadores. Algo que era comum ser aplicado por alguns times apenas em vitórias, finalizações e gols, passou a ser aplicado com outras variáveis pela ótica de Graham, que usava outras variáveis, passadas para os jogadores de maneira resumida, a fim de que eles soubessem seu objetivo para obter uma melhora individual, e, conseqüentemente, melhorar a desenvoltura do time.

Os Estados Unidos são um dos maiores exemplos do uso da Estatística no esporte, segundo o artigo do site “Quero Apostar”, ela é utilizada nos principais esportes do país, como na NBA (National Basketball Association), liga de Basquete dos EUA, no uso da eficiência ofensiva e defensiva das equipes, além da utilização do “pace” que é uma forma de analisar a movimentação de ataque dos times, se é rápida ou lenta, sendo assim, gerando estratégias a partir desse ponto. A NFL (National Football League), liga de futebol americano, também usa bastante a ideia de dados sobre ataque e defesa dos times e analisa em específico o “turnover”, termo utilizado quando um time tem a posse da bola e perde para o time adversário, lance que muitas vezes decide a partida. Outra ligas, como a MLB, de basebol, oferece um acervo gigante de dados para serem analisados. Já a NHL (National Hockey League), liga de hockey, é a que oferece menor aplicação da estatística, mas, ainda assim, utiliza a análise de ataque e defesa dos times. Outros esportes como a Fórmula 1, o Tênis e o Atletismo também utilizam a estatística em seu processo de estudos e tomadas de decisões. Vale lembrar que, além da estatística, deve-se considerar as condições psicológicas dos jogadores, que influenciam diretamente no desempenho desses atletas.

Segundo a matéria “Os números que mudam o jogo: a influência dos dados no esporte”, no site Liga Insights, a área de sports analytics deverá alcançar até o ano de 2025 um a marca de US\$ 4,6 bilhões como valor de mercado, que tem como uma das suas principais áreas o “Scout” de jogadores, utilizado para análise do desempenho

que auxilia no processo de seleção para algum time, com tudo desenvolvido em aplicativos. Também existe o “Startups”, uma ferramenta de captação de dados dos jogadores, sendo eles dados físicos, técnicos, táticos e psicológicos, que está em processo de aplicação no Brasil pela Startup Soul Brasil.

A empresa de startups paulistana, iSPORTiSTiCS, trabalha com dados e marketing esportivo, área que crescem no Brasil, aliando a tecnologia e a estatística para criação de vídeos para divulgação de eventos esportivos. Porém, o processo de inserção da tecnologia no esporte no Brasil é difícil, devido à cultura existente de não aceitar mudanças que possam modificar a modalidade. Vinicius Gholmie, fundador da iSPORTiSTiCS, em entrevista para a Liga Insights, em 2019, disse “O baixo conhecimento dos envolvidos nas decisões ainda é um entrave. Tendo a acreditar pouco em mudanças contundentes em um segmento, se essas não forem conduzidas pelos líderes. No mínimo, elas serão muito demoradas, dolorosas. Ou seja, quem tem que difundir a inovação são os principais agentes do mercado esportivo: grandes times, ligas, marcas, federações, os maiores grupos de mídia. A partir deles fica muito mais fácil cobrir todo o mercado e trazer uma mudança contundente”.

Esse trabalho apresentará a estatística de maneira simples, com a apresentação da comparação do desempenho de times de futebol, por meio de coleta de dados, organização das informações e apresentação em tabelas, gráficos ou medidas, com o uso software Excel como ferramenta. Essa ideia está fundamentada nas palavras de Silva ([21], p.3, apud Sampaio e Leite, 2008, p. 19), para o qual o papel do professor deverá sofrer transformações, devido a todo o ritmo quem vem tomando, com acesso à muita informação, o que faz com que o professor precise analisar e compreender a realidade circundante, a fim de buscar o domínio das tecnologias e, assim, não seja dominado por elas. Com a união desse pensamento, de apresentação da estatística como uma ferramenta de comparação junto à tecnologia, afirmamos que o nosso objetivo será de ensinar a estatística por meio de uma comparação entre times de futebol em várias edições do campeonato, com a utilização, portanto, de diferentes gráficos, medidas de posição e medidas de tendência central. Isso poderá ser aplicado facilmente dentro de sala de aula com os alunos, e sofrer modificações de acordo com o desejo do professor.

Além desse capítulo de introdução, teremos outros quatro. No segundo, abordaremos as noções de estatística, por meio de conceitos fundamentais, de tabelas, medidas de posição e medidas de dispersão, traremos vários conceitos e exemplos, com objetivo do leitor ter noção do que está sendo trabalhado. No terceiro capítulo, será abordada a comparação estatística, por meio das ferramentas oferecidas no capítulo anterior, nele foi feita uma coleta de dados dos campeonatos brasileiros e organizados para que fosse possível fazer uma comparação. O quarto, será apresentado a utilização do Excel, assim como o passo a passo do que é necessário para uma comparação estatística, que foi colocado como capítulo pela sua importância no desenvolvimento do trabalho, mostrando ferramentas básicas para o desenvolvimento. No quinto capítulo teremos a conclusão do trabalho.

Capítulo 2

Conceitos Fundamentais da Estatística

Neste capítulo, apresentaremos os conceitos estatísticos e ilustraremos exemplificando com o objetivo de servir como referencial para os capítulos seguintes, nos quais compararemos clubes em diferentes edições do Campeonato Brasileiro de Futebol. A teoria apresentada é baseada nos livros de “Probabilidade e Estatística” de Correa [4], “Aprender fazendo estatística ” de Guedes [12], “Estatística básica” de Bussab [2], “Estatística Básica” de Ferrari [9], “Compreensões de alunos e professores sobre Média Aritmética” de Marques [16], “Probabilidade e Estística aplicada ao esporte” de Ribeiro [19] e “Matemática e realidades: interdisciplinar” de Salsa [20].

2.1 Noções básicas

Para Correa [4], a “Estatística é uma área da Matemática que oferece métodos de coleta, organização, descrição, análise e interpretação de dados, viabilizando a utilização destes para uma tomada de decisão”. Com tais considerações, apresentaremos as noções e conceitos iniciais da estatística, que servem para seu ensino, podendo assim, mostrar os exemplos dentro do contexto em que a dissertação está envolvida.

2.1.1 População

É um conjunto completo de todos os elementos a serem estudados. Por exemplo, os estudantes de uma escola, indivíduos de uma cidade e, para nossa situação, clubes de um campeonato brasileiro de futebol da primeira divisão (Série A).

2.1.2 Amostra

É um subconjunto de elementos da população, ou seja, são elementos selecionados do conjunto maior. Por exemplo, os estudantes de uma turma de uma determinada escola, população de um bairro de uma cidade, os quatro primeiros colocados da primeira divisão do campeonato brasileiro de futebol da primeira divisão (Série A) da edição 2020.

2.1.3 Censo

É uma coleção de dados referentes aos elementos da população. Tal como as notas dos estudantes da escola, local de moradia das pessoas de uma cidade e região do país dos clubes da série A.

2.1.4 Parâmetro

É uma medida numérica fixa que geralmente essa é desconhecida. Para exemplificar, temos a média das notas dos estudantes da escola, moda das idades das pessoas de uma cidade e desvio padrão do número de pontos dos times campeões da série A.

2.1.5 Variável

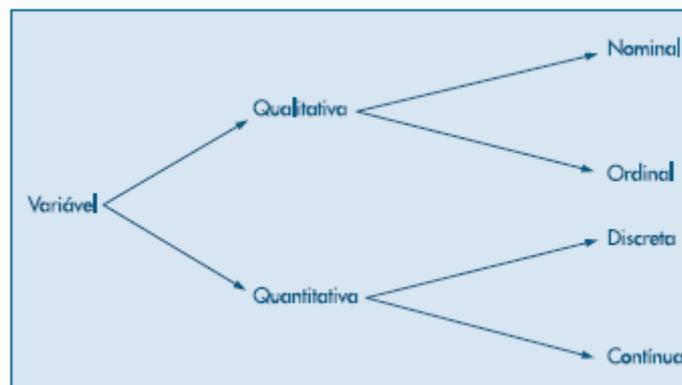
As variáveis dentro da pesquisa é o que podemos chamar de característica, isto é, o número de familiares na casa dos estudantes de uma escola, estado civil das pessoas de uma cidade ou número de finalizações de um time da série A do campeonato brasileiro. As variáveis são divididas em 2 tipos:

Variável qualitativa é aquela que é relacionada a uma característica de qualidade ou atributo, como nível de escolaridade, gênero sexual etc.

Variável quantitativa é a variável relacionada a uma característica que pode ser mensurada ou medida, tal como número de gols, número de finalizações etc.

Tanto as variáveis qualitativas, quanto quantitativas podem ser, ainda, classificadas em outros dois subgrupos, cada uma. A qualitativa é dividida em **ordinal** e **nominal**, já a quantitativa é dividida em **discreta** e **contínua**. Observemos o quadro na Figura 2.1.

Figura 2.1: Classificação das variáveis.



Fonte: Bussab, p.10

A **variável qualitativa nominal** é aquela que não existe ordenação em suas características, por exemplo, cidade de nascimento dos jogadores de um time de futebol: Natal, Recife, Rio de Janeiro, São Paulo etc.

Já a **variável qualitativa ordinal** existe uma ordenação em seus resultados, por exemplo, o grau de escolaridade dos jogadores do Flamengo, ensino fundamental, ensino médio, ensino superior etc.

Quando falamos da **variável quantitativa discreta**, nos referimos às características cujo os valores possíveis formam um conjunto finito, que resultam em uma contagem. Um exemplo é o número de gols marcados por uma equipe durante o campeonato nacional, teremos valores dentre os números $(0, 1, 2, 3, \dots)$.

A **variável quantitativa contínua** relaciona os valores contidos nos números reais, que podemos medir. Por exemplo, média de gols por jogo dos jogadores do Palmeiras, que sabemos que será uma divisão entre o número de gols pelo número de jogos. Pode ser também o peso dos jogadores ou altura.

Entender as variáveis e como elas podem gerar resultados é importante para colher os dados certos, com objetivo de ter uma pesquisa satisfatória, para obter uma comparação correta, a título de exemplo, a comparação dos rendimentos de alguns clubes e o que levou um time a ter êxito ou não.

2.1.6 Tabela

Segundo Correa ([4], p. 16), tabela é um quadro que contém os resumos das características com objetivo de compará-los. Toda tabela é composta por corpo, Cabeçalho, Coluna indicadora, Linhas, Casa ou célula, Título e Fonte.

Iremos definir todos os elementos de uma tabela e usar como exemplo a Tabela 2.1, que é a de classificação final do campeonato brasileiro de futebol de 2019.

Tabela 2.1: Classificação do campeonato brasileiro de futebol de 2019.

Ps	TIME	P	J	V	E	D	GP	GC	SG
1	Flamengo	90	38	28	6	4	86	37	49
2	Santos	74	38	22	8	8	60	33	27
3	Palmeiras	74	38	21	11	6	61	32	29
4	Grêmio	65	38	19	8	11	64	39	25
5	Athletico	64	38	18	10	10	51	32	19
6	São Paulo	63	38	17	12	9	39	30	9
7	Internacional	57	38	16	9	13	44	39	5
8	Corinthians	56	38	14	14	10	42	34	8
9	Fortaleza	53	38	15	8	15	50	49	1
10	Goiás	52	38	15	7	16	46	64	-18
11	Bahia	49	38	12	13	13	44	43	1
12	Vasco	49	38	12	13	13	39	45	-6
13	Atlético Mineiro	48	38	13	9	16	45	49	-4
14	Fluminense	46	38	12	10	16	38	46	-8
15	Botafogo	43	38	13	4	21	31	45	-14
16	Ceará	39	38	10	9	19	36	41	-5
17	Cruzeiro	36	38	7	15	16	27	46	-19
18	CSA	32	38	8	8	22	24	58	-34
19	Chapecoense	32	38	7	11	20	31	52	-21
20	Avaí	20	38	3	11	24	18	62	-44

Fonte: <https://www.cbf.com.br/futebol-brasileiro/competicoes/campeonato-brasileiro-serie-a/2019>

Corpo é todo conjunto de linhas e colunas que contém as informações das variáveis. Observemos que na Tabela 2.1 corpo é tudo que está dentro das linhas e colunas.

Cabeçalho é a parte superior da tabela, que especifica o conteúdo da coluna, por exemplo, o ranking (R), os times e as quantidades de pontos (P), jogos (J), vitórias (V), empates (E), derrotas (D), gols pró (GP), gols contra (GC) e saldo de gols (SG).

Coluna indicadora é a parte da tabela que especifica o conteúdo da linha, por exemplo, o posicionamento do time, que indica o 1º, 2º, etc.

Linhas são as retas no sentido horizontal, por exemplo as linhas que contém as informações de cada clube.

Casa ou célula é o espaço destinado a um número. É onde está cada informação individual do time, a interseção entre as linhas e colunas, em forma de retângulo.

Título é o conjunto de informações da tabela, quando foi e onde foi. No caso da nossa tabela é o Campeonato Brasileiro de futebol de 2019.

Fonte é a referência de onde se obteve os dados da tabela, na nossa situação o site da CBF, que é a Confederação Brasileira de Futebol.

2.2 Distribuição de frequência e intervalo de classe

Para entender a variável e saber como funciona o comportamento dos dados existe a distribuição de frequência, que segundo Bussab ([2], p.11), serve para organizar valores, onde cada entrada teria uma frequência. As informações podem ser colocadas em ordem, que chamamos de rol. As variáveis quantitativas e qualitativas podem ser colocadas em rol, a quantitativa geralmente é ordem crescente ou decrescente e a qualitativa é geralmente colocada em rol quando tem-se uma ordem alfabética ou atemporal.

As frequências são classificadas em absolutas e relativas. A **frequência absoluta**, representada por f , é o número ocorrências da variável, já a **frequência relativa**, representada por f_r , é a razão entre a frequência absoluta e o número de elementos da população, ou seja, $f_r = \frac{f}{n}$. A relativa pode ser representada como fração, decimal e porcentagem.

Na Tabela 3.2, observamos que a frequência é absoluta, que é o número de gols marcado por cada time, ordenado de maneira decrescente, do time que marcou mais gols para o que marcou menos gols. A frequência é o número de gols marcado por cada time, por exemplo, o Flamengo marcou 86 gols, ou seja, o gol ocorreu em uma frequência de 86 vezes. Em relação a f_r , sabemos que o número de elementos é o total de gols marcados no campeonato, que foi de 876, calculamos que a f_r do Flamengo

foi de $f_r = 86/876 = 0,0982$, aproximadamente.

Tabela 2.2: Gols marcados pelos times da série A de 2019.

R	TIME	GP
1	Flamengo	86
2	Grêmio	64
3	Palmeiras	61
4	Santos	60
5	Athletico	51
6	Fortaleza	50
7	Goiás	46
8	Atlético Mineiro	45
9	Bahia	44
10	Internacional	44
11	Corinthians	42
12	Vasco	39
13	São Paulo	39
14	Fluminense	38
15	Ceará	36
16	Botafogo	31
17	Chapecoense	31
18	Cruzeiro	27
19	CSA	24
20	Avaí	18

Fonte: <https://www.srgool.com.br/classificacao/Brasileirao/Serie-A/2019?a=todosclassificacao>

Quando falamos de intervalos, nos referimos ao conjunto dos números reais, que é o conjunto de números limitados por dois extremos, na estatística temos o chamado intervalos de classe, que segundo Ferrari ([9], p.14) são os intervalos de variações das variáveis, representada por $i = 1, 2, 3, \dots, k$, com k o total de intervalos e n o número de valores totais. Podemos representar por L_i o limite superior da i -ésima classe e por l_i o limite inferior da i -ésima classe, e com isso definir $h_i = L_i - l_i$ como o tamanho do intervalo de classe.

A frequência também pode ser dita acumulada, representada por F_i , e definida como a soma das frequências absolutas de todas as classes com intervalos inferiores à classe dada. Pode ser calculada por $F_j = \sum_{i=1}^j F_i = F_1 + F_2 + \dots + F_j$.

2.3 Gráficos

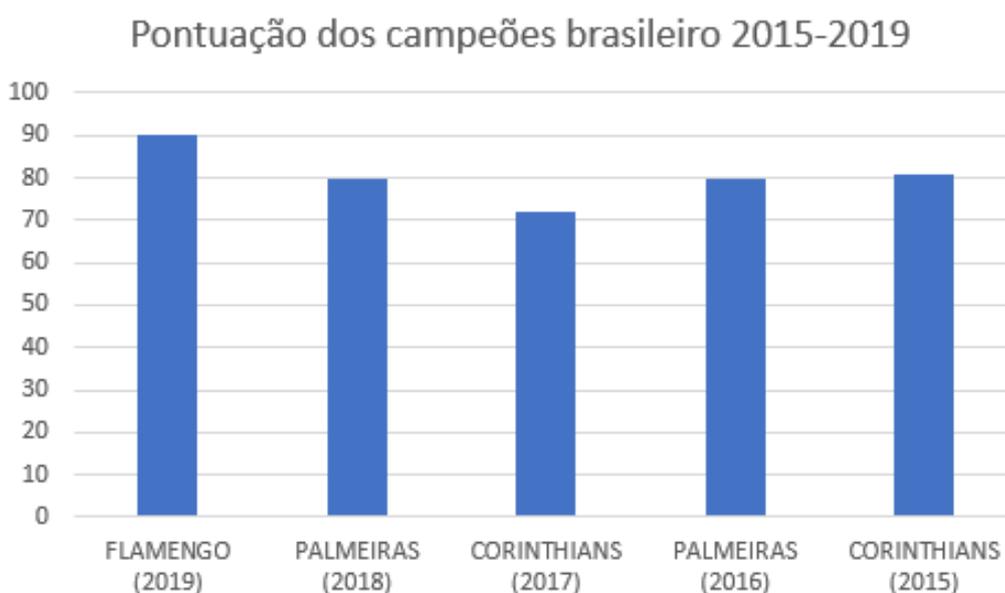
Os gráficos, segundo Bussab ([2], p. 15), são representações das distribuições das variáveis, que trazem para os leitores uma interpretação mais rápida, dinâmica e abrangente da pesquisa. Existem vários tipos de gráficos que podemos utilizar de acordo com tipos de variáveis, sendo elas variáveis qualitativas ou quantitativas. Mostraremos alguns gráficos que são usuais no futebol.

2.3.1 Gráfico de barras ou colunas

O gráfico de barras é utilizado para representações das variáveis qualitativas e quantitativas, feito para resumir as informações por meio de retângulos (barras), na vertical ou horizontal. Um ponto forte é a fácil comparação entre as informações, apenas com a visualização.

Exemplo 2.3.1. *A Figura 2.2 é um exemplo de como pode ser utilizado no futebol. Nesse caso, temos a pontuação dos últimos 5 campeões brasileiros, representados por um gráfico de barras (colunas).*

Figura 2.2: Gráfico da pontuação dos campeões brasileiros de 2015 até 2019



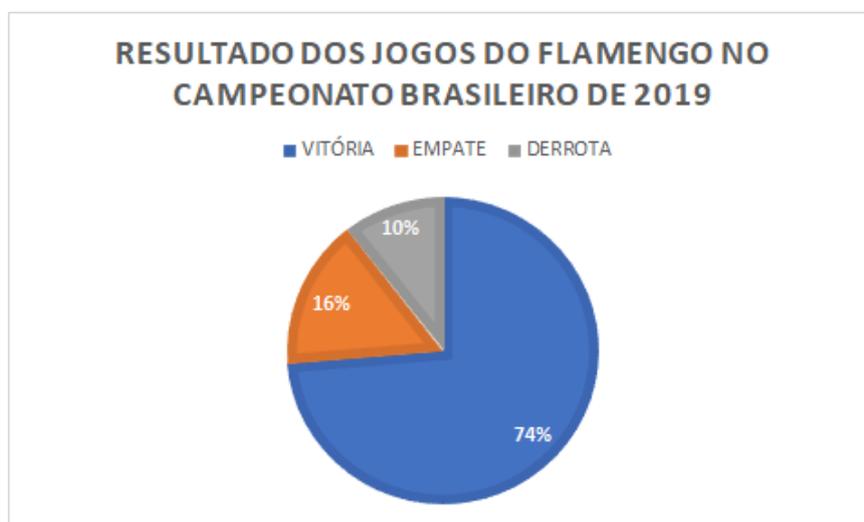
Fonte: CBF

2.3.2 Gráfico de setores

É um gráfico bastante utilizado com as variáveis qualitativas, representado em uma composição circular, em formato de “pizza”. Seu uso é indicado para situações que usem porcentagem, pois é mais fácil de ver quanto de um todo aquela parte representa.

Exemplo 2.3.2. A Figura 2.3 mostra os resultados, por meio do gráfico de setores, do Flamengo no campeonato de 2019, sabemos que o total é de 38 jogos, representado o 100% e cada variável são mostrados em número de vitórias, empates e derrotas, exposta em porcentagem, em relação ao total de jogos.

Figura 2.3: Gráfico do resultado do Flamengo no campeonato brasileiro de 2019



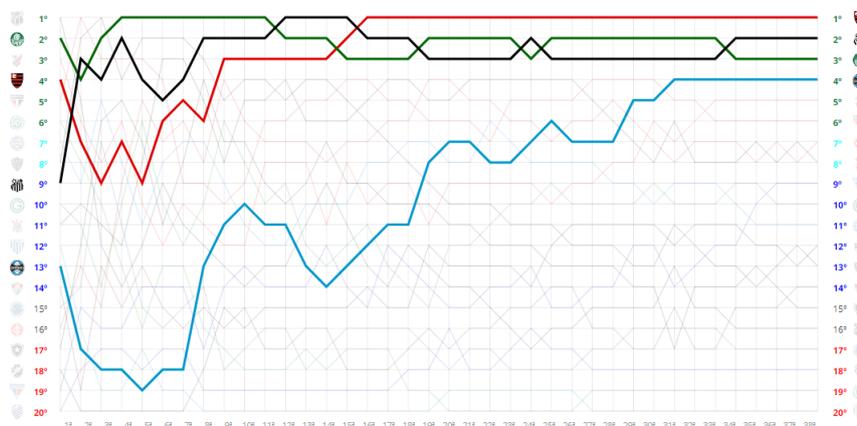
Fonte: CBF

2.3.3 Gráficos de segmentos ou linha

Conhecido como gráfico de linha ou segmento, é geralmente utilizado em variáveis quantitativas discretas, que mostram as evoluções das variáveis pelas suas frequências absolutas e acumulativas.

Exemplo 2.3.3. A Figura 2.4 mostra o desempenho dos 20 clubes no campeonato brasileiro de 2019, durante as 38 rodadas, por meio do gráfico de segmentos.

Figura 2.4: Gráfico do desempenho dos times do campeonato brasileiro de 2019, destaque para os 4 primeiros colocados.



Fonte: CBF

2.3.4 Gráfico em área

Uma outra representação para os gráficos de colunas ou linhas pode ser o gráfico em área, pois ele tem a vantagem de trazer uma percepção maior das mudanças da variável, que chama atenção para o valor total da variável. O componente da área representa o total acumulado da variável, no caso é a pontuação, que segundo o site HighBond [26], é um tipo de gráfico de área é chamado de padrão, apenas os valores totais.

Exemplo 2.3.4. A figura 2.5 mostra o desempenho do Vasco nas oito primeiras rodadas do campeonato, com suas pontuações acumuladas.

Figura 2.5: Pontuação acumulada do Vasco nas 8 primeiras rodadas

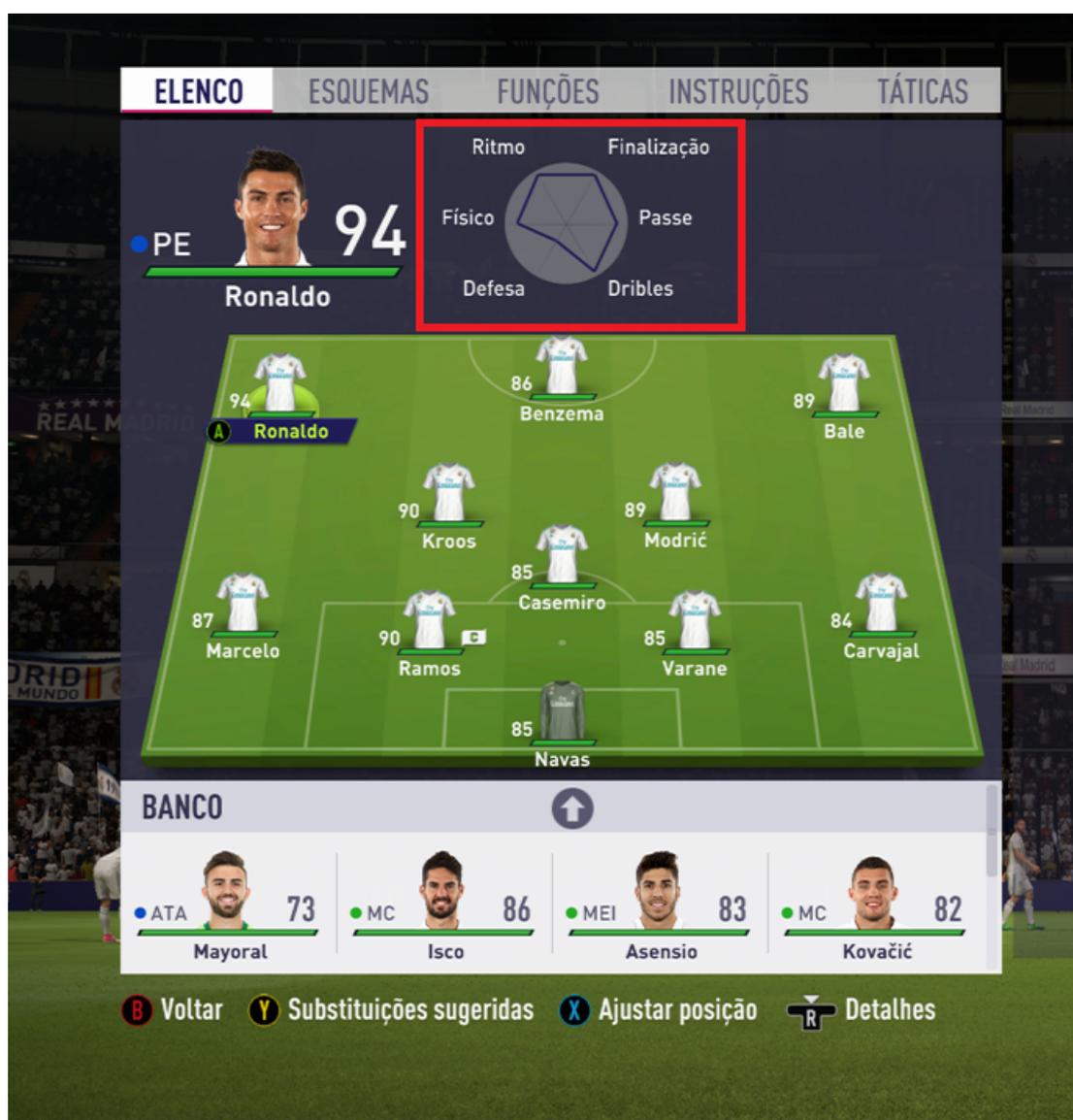


2.3.5 Gráfico em radar

Outro gráfico que também pode ser utilizado com as informações dos gráficos em colunas ou linhas é o em radar. Tem como vantagem obter uma outra percepção entre comparações de variáveis.

Exemplo 2.3.5. A Figura 2.6 nos mostra uma das utilizações do gráfico de radar, em jogos de vídeo game, como no FIFA 18. Na figura encontramos o gráfico em radar que contempla as atribuições do jogador Cristiano Ronaldo, destacado em vermelho.

Figura 2.6: Gráfico em radar de Cristiano Ronaldo no FIFA 18



Fonte: Acesso em: <https://indicadoresportivo.com/grafico-radar-no-fifa-18/>

2.4 Medidas de posição

As medidas de posição, ou medidas de tendência central, são valores que representam todo o conjunto numérico, que busca entender onde está a maior concentração de valores e visualiza o seu posicionamento. O principal objetivo é fazer um confronto desses números, para comparar as variáveis. As medidas de posição mais utilizadas são a média aritmética, mediana e moda. Além das mais comuns, iremos apresentar a média geométrica, média harmônica e separatrizes, que são valores que ocuparam lugares que separam a quantidade de elementos do conjunto.

2.4.1 Média Aritmética

A média aritmética é utilizada em várias situações do nosso cotidiano, nas escolas, nos trabalhos, no meio econômico e também no contexto esportivo. Antes de definir o que é a média aritmética e como calculá-la, temos que entender o objetivo dela. Segundo Marques ([16], p. 727, apud Carzola e Santana, 2006, p.18) a média aritmética resume a variável, ou o conjunto numérico, em apenas um valor, sendo aquele valor que irá “representar” o conjunto. Entendendo o objetivo da média, podemos definir a **média aritmética** como uma razão ou quociente, entre a soma de todos os valores da variável pela quantidade de valores, que é expressa pela fórmula a seguir, segundo Correa [4]:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}.$$

Sendo \bar{x} a representação da média, x_i os valores, i o índice (a posição do valores), que varia do primeiro ao último e n o número de termos.

Exemplo 2.4.1. *Para calcular a média de gols de Claudinho do Bragantino no Brasileiro de 2020, basta somar todos os seus gols em cada jogo e dividir pelo total de jogos. Claudinho marcou 18 gols em 35 jogos, logo sua média é de $\frac{18}{35} = 0,51$ gols por jogo, ou seja, significa que ele tem aproximadamente 1 gol a cada 2 jogos, sendo essa, a representação da sua média.*

2.4.2 Média Ponderada

A média ponderada tem o significado análogo com a média aritmética, cujo o diferencial é a repetição dos valores das variáveis, que é popularmente chamado por “pesos”, ou seja, é uma forma de ponderação. O cálculo da média ponderada é da seguinte forma, segundo Correa [4]:

$$\frac{\sum_{i=1}^n (p_i \cdot x_i)}{\sum_{i=1}^n p_i}.$$

Com x_i expressando os valores, p_i os pesos (frequências) e i o índice, no qual varia 1 até n (último termo). Desenvolvendo a expressão, ficamos com:

$$\frac{p_1 \cdot x_1 + p_2 \cdot x_2 + p_3 \cdot x_3 + \dots + p_n \cdot x_n}{p_1 + p_2 + p_3 + \dots + p_n}.$$

Exemplo 2.4.2. *Para média de gols de Gabriel, artilheiro do Flamengo no campeonato Brasileiro de 2019, nas 17 primeiras partidas. Marcou 2 gols em 5 oportunidades, 1 gol em 9 oportunidades e nenhum gol em outras 3 oportunidades, logo sua média, pode ser calculada como uma média ponderada, que utiliza a quantidade de partidas, como um “peso”.*

$$\frac{5 \cdot 2 + 9 \cdot 1 + 3 \cdot 0}{5 + 9 + 3} = \frac{10 + 9 + 0}{17} = \frac{19}{17} = 1,11.$$

2.4.3 Média Geométrica

A média geométrica é indicada em situações em envolva porcentagens. O cálculo é feito pela raiz n -ésima da multiplicação dos n valores. A média é representada por \bar{x}_g ou M_g e calculada pela fórmula, segundo Correa [4]:

$$\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}$$

Onde X_i são os termos que vão variar de 1 até n (último), lembrando que n também é o índice da raiz.

Exemplo 2.4.3. *Imaginemos que um time, em seus 3 primeiros jogos, tenha um percentual de finalizações no gol, em relação ao total de chutes, de 80% no 1º jogo, 75% no segundo jogo e 88% no 3º jogo. Sabendo que a quantidade de finalizações e o total de gols não são conhecidas, devemos utilizar a média geométrica para calcular, pois a aritmética ou ponderada não seria indicada.*

$$\sqrt[3]{0,8 * 0,75 * 0,88} = \sqrt[3]{0,528} = 0,8082,$$

ou seja, 80,82%. Sendo essa, a média de acertos no alvo, nas 3 primeiras partidas.

2.4.4 Média Harmônica

Tem como principal característica a aplicação em valores em forma de fração, segundo Ribeiro ([19], p.46), é quando os valores estão caracterizados de maneira inversamente proporcional, representada por \bar{x}_h ou M_h . Ela será calculada da seguinte maneira:

$$\bar{x}_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}},$$

onde, x_1, x_2, \dots, x_n são os valores e n a quantidade de valores.

Uma aplicação importante da média harmônica é a média das velocidades, que não pode ser confundida com a velocidade média, que segundo Bernardo ([1], p.58), é a razão entre o espaço percorrido e o intervalo de tempo gasto, geralmente encontrada com o conhecimento dos valores do espaço percorrido e o tempo passado. Já a média das velocidades, que é representada pela média harmônica, é utilizada com o conhecimento das velocidades e com o objetivo de encontrar a média, que por sua vez está relacionado à ideia de velocidade instantânea.

Exemplo 2.4.4. *Em um jogo do PSG, Mbappé atingiu a marca de 36 km/h em um ataque do clube e uma marca de 30 km/h, na recomposição para defender, na sequência do lance. Para calcular a média da velocidades dele utilizaremos a média harmônica:*

$$\frac{2}{\frac{1}{36} + \frac{1}{30}} = \frac{2}{\frac{5+6}{180}} = \frac{2}{\frac{11}{180}} = \frac{360}{11} = 32,72 \text{ km/h.}$$

2.4.5 Mediana

A mediana faz parte de um grupo chamado medidas separatrizes, que dividem a distribuição em grupos com mesma quantidade de elementos, pode ser dividido em dois grupos, quatro grupos, etc. Existem outros tipos de separatrizes, como os Quartis, dividem a distribuição em 4 partes iguais, com 1º Quartil que separa o conjunto em 25% dos valores menores e 75% dos valores maiores, o 2º Quartil tem 50% dos valores menores e 50% dos valores maiores e o 3º Quartil tem 75% dos valores menores e 25% dos valores maiores. Existem outras separatrizes como os Decis, Percentis etc., mas definiremos apenas a mediana, que são separatrizes que decompõem em 50% valores menores e 50% valores maiores que a mediana.

Quando os dados são colocados em tabelas, chamamos de dados tabelados. Caso contrário, dizemos que são não tabelados, no caso dos números estarem apenas listados. Em ambos os casos, quando são colocados em ordem, chamamos de rol.

- **Para dados não tabelados**

Para encontrar a posição (valores) dessa mediana para dados não tabelados, existe duas situações: uma série com um número par de elementos e uma série com um número ímpar de elementos. Vale lembrar que é importante esses elementos estarem em rol.

Com um **número ímpar** de elementos, a mediana vai ser o valor do meio. Sendo n o número de itens, a posição da mediana será dada por:

$$E_{Md} = \frac{n + 1}{2}.$$

Onde E_{Md} é o elemento mediano.

Já em um caso com **número par** de elementos, a mediana vai ser a média aritmética dos dois valores centrais. Para achar a posição dos dois valores usaremos:

$$E_1 = \frac{n}{2}$$

e

$$E_2 = \frac{n + 2}{2}$$

Com E_1 e E_2 os elementos mediano. Dizemos que a mediana é a média deles, ou seja, $M_d = \frac{E_1 + E_2}{2}$

Exemplo 2.4.5. *Os jogadores do Flamengo com mais de 1,80 m são Thuler e Vitor Gabriel, com 1,85m e 1,84, respectivamente, já Vasco tem-se Werley com 1,84 m, Leandro Castan com 1,86 m, Danilo com 1,82 m, Lucas Mineiro com 1,88 m, Marrony com 1,88 m e Maxi Lópes com 1,85 m.*

Colocado em rol (ordem) os valores dos jogadores com mais de 1,80 m do Flamengo e Vasco, temos:

1,82m; 1,84m; 1,84m; 1,85m; 1,85m; 1,86m; 1,88m; 1,88m.

Com isso, como temos 8 itens (número par), a mediana será a média dos 2 valores do médio, no caso na posição $\frac{8}{2} = 4$ e 5, ou seja, $\frac{1,85 + 1,85}{2}$, resultando em 1,85 m.

Há outras maneiras de calcular a mediana, que é no caso dos dados aparecerem agrupados sem intervalo de classe ou agrupados em intervalos de classe. Vamos observar cada uma dessas outras formas.

- **Para dados tabelados e não agrupados em classe**

Calcularemos a mediana encontrando a posição dela, como fazemos sem dados tabelados, com o uso da ideia de ser par ou ímpar. Em sequência, devemos encontrar o valor numérico correspondente a posição, por meio da frequência acumulada.

Exemplo 2.4.6. *Vamos observar a tabela a seguir com a quantidade de gols do Palmeiras, por jogo, no campeonato brasileiro da série A de 2020.*

Tabela 2.3: Gols por jogo do Palmeiras na série A 2020

Número de gols	Frequência	Frequência acumulada
0	7	7
1	18	25
2	7	32
3	5	37
4	1	38

Fonte: <https://www.espn.com.br/futebol/time/resultados/id/2029/palmeiras>

A quantidade de partidas foram 38, ou seja, a soma das frequências resulta em 38. Com isso temos dois valores para serem encontrados são da posição: $38/2 = 19^\circ$ e 20° . Pela frequência acumulada, ambos os valores pertencem ao número de 1 gol por jogo, pois são maiores que 7 e menores que 25. Fazendo a média aritmética entre 1 e 1, vamos obter $M_d = 1$ como o valor da mediana.

- **Para dados tabelados e agrupados em classe**

Os valores que serão tabelados e com intervalo de classe precisam que encontremos a classe mediana pela fórmula, segundo Correa [4]:

$$M_d = l_i + h \cdot \frac{\sum_{i=1}^n \frac{f_i}{2} - F_{ant}}{f_{Md}}$$

Onde l_i é o limite inferior da classe; F_{ant} é a frequência acumulada anterior ao intervalo; f_{Md} é a frequência simples da classe; h é a amplitude do intervalo; e $\sum_{i=1}^n \frac{f_i}{2}$ é o 1º passo, que iremos encontrar a posição da mediana.

Exemplo 2.4.7. *Como exemplo tomemos as alturas dos jogadores do Atlético Mineiro.*

Tabela 2.4: Altura dos jogadores do Atlético MG.

Altura (m)	Frequência	Frequência acumulada
1,50 † 1,60	2	2
1,60 † 1,70	7	9
1,70 † 1,80	7	16
1,80 † 1,90	3	19
1,90 † 2,00	1	20

O 1º passo é encontramos a posição da mediana, ou seja, $\frac{20}{2} = 10^\circ$ termo pertence a classe de 1,70m à 1,80m.

Para encontrar a mediana vamos usar a fórmula:

$$M_d = l_i + h \cdot \frac{\sum_{i=1}^n \frac{f_i}{2} - F_{ant}}{f_{Md}}.$$

$l_i = 1,70$ é o limite inferior da classe;

$n = \sum_{i=1}^n f_i = 10 = 20$ é o número de elementos;

$f = 7$ é a frequência da classe da mediana;

$h = 0,1$ é a amplitude da classe da mediana;

$F_{ant} = 9$ é a frequência acumulada da classe anterior.

Logo $M_e = 1,70 + 0,1 \cdot \frac{10-9}{1,70} = 1,76m$, *aproximadamente*.

2.4.6 Moda

Segundo Bussab ([2], p.35), definimos como moda valor que tem a maior frequência, ou maior aparição no conjunto de valores e podemos representar a moda por M_o .

Salsa ([20], p.14) afirma que há nomenclaturas que dependem das quantidade de modas que estão contidas em um conjunto. Quando tem-se 2 valores com a maior frequência, dizemos que o conjunto é bimodal, no caso em que temos 3 valores com a maior frequência, chamamos de trimodal, com 4 ou mais, usamos o plurimodal. A ausência de moda e para casos com muitos números com a maior frequência usamos o termo amodal.

Exemplo 2.4.8. *Observemos na figura 2.7, que contém a pontuação, a quantidade de jogos e o número de vitória dos times da Série A de 2019. Queremos calcular a moda*

do número de vitórias, para isso basta observar o número que aparecer com a maior frequência, ou seja, qual a quantidade de vitórias que aparecem mais vezes. Vejamos que o 12 aparece em 3 oportunidades, com o Bahia, Vasco e Fluminense, sendo 12 vitórias o número que mais se repetiu ao final do campeonato.

Figura 2.7: Número de vitória dos times da série A de 2019.

1º	0	 Flamengo - RJ	90	38	28
2º	0	 Santos - SP	74	38	22
3º	0	 Palmeiras - SP	74	38	21
4º	0	 Grêmio - RS	65	38	19
5º	0	 Athletico Paranaense - PR	64	38	18
6º	0	 São Paulo - SP	63	38	17
7º	+1	 Internacional - RS	57	38	16
8º	-1	 Corinthians - SP	56	38	14
9º	0	 Fortaleza - CE	53	38	15
10º	0	 Goiás - GO	52	38	15
11º	0	 Bahia - BA	49	38	12
12º	+1	 Vasco da Gama - RJ	49	38	12
13º	-1	 Atlético - MG	48	38	13
14º	0	 Fluminense - RJ	46	38	12
15º	0	 Botafogo - RJ	43	38	13
16º	0	 Ceará - CE	39	38	10
17º	0	 Cruzeiro - MG	36	38	7
18º	0	 Csa - AL	32	38	8
19º	0	 Chapecoense - SC	32	38	7
20º	0	 Avaí - SC	20	38	3

Fonte: CBF

2.5 Medidas de dispersão

Para Bonita ([26], p.68), as medidas de dispersão servem para resumir dados e mostrar o quão são diferentes os elementos do conjunto. Podemos dizer que o grau de variação dos elementos de um conjunto, aplicadas em situações que tem como característica a semelhança e a variabilidade, auxiliando, assim, as medidas de tendência central a descrever o conjunto e os dados de uma maneira mais precisa.

Imaginemos uma situação hipotética em que o time A queira contratar um entre os três jogadores, Gil, Iago e Cleiton. Tem-se apenas as informações dos gols que eles marcaram nos anos de 2018, 2019 e 2020, nessa ordem:

$$\text{Iago} = \{ 20, 20, 20 \} \Rightarrow M = \frac{20+20+20}{3} = 20 \text{ gols}$$

$$\text{Gil} = \{ 20, 15, 25 \} \Rightarrow M = \frac{20+15+25}{3} = 20 \text{ gols}$$

$$\text{Cleiton} = \{ 30, 25, 5 \} \Rightarrow M = \frac{30+25+5}{3} = 20 \text{ gols}$$

A medida de tendência central não é suficiente para fazer essa contratação, pois todos os jogadores têm a mesma média de gols. Para isso precisamos trabalhar com as medidas de dispersão, pois elas irão apresentar a dispersão dos valores. Iago, por exemplo, a média condiz com os valores, todos os três anos ele marcou 20 gols, já no caso de Gil, ficou com apenas uma diferença de 5 gols para mais e para menos em 2 anos, sendo que um deles ele marcou 20 gols. Por fim o Cleiton, que teve uma distância bem maior dos valores para a média, que marcou 30 e 25 gols nos dois primeiros anos, os quais compensam o baixo rendimento no último ano. É fato que alguns fatores físicos ou psicológicos podem ter influenciados os rendimentos nos anos com menos gols, como por exemplo uma contusão ou problemas pessoais, estamos levando em conta apenas os números, mas também poderíamos ter levado em conta esses outros fatores.

Essa situação nos mostra ideia de regularidade, que é bastante utilizada no futebol e em outros esportes, na qual o jogador com o menor desvio é mais regular, pois tem seus valores mais próximos da média. Analogamente, o jogador com o maior desvio é menos regular. Por esse motivo, o Iago deveria ser o contratado, pois ele manteve

uma maior regularidade nos 3 últimos anos, diferentemente de Gil e Cleiton, que acabaram oscilando, ou seja, com uma menor regularidade. Para aprofundarmos o estudo das medidas de dispersão.

2.5.1 Amplitude total

A Amplitude total é definida como a diferença entre o maior e o menor elemento de um conjunto numérico, ou seja, $A_t = (\text{maior valor}) - (\text{menor valor})$.

Exemplo 2.5.1. *Cleiton fez 30 gols em 2018, 25 gols em 2019 e 5 gols em 2020, podemos afirmar que a amplitude dos gols dele nos três últimos anos foi de: $A_t = 30 - 5 = 25$ gols.*

2.5.2 Desvio Médio

Apresentaremos apenas o desvio médio amostral, que é calculado para informar dispersão da média e dos valores do conjunto, cujo é calculado pela média da diferença entre a média do conjunto e cada elemento, dado pela fórmula:

$$D_m = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

Sendo x_i cada valor e \bar{x} a média do conjunto. Caso os dados sejam dados por uma distribuição de frequência usamos:

$$D_m = \frac{\sum_{i=1}^n |x_i - \bar{x}| F_i}{n}$$

Com F_i sendo a frequência do conjunto.

Exemplo 2.5.2. *Gil marcou 20, 15 e 25 gols nos anos de 2018, 2019 e 2020, respectivamente, enquanto Cleiton marcou 30, 25 e 5. Vamos calcular o desvio médio de*

seus valores, sabendo que a média de ambos é de 20.

$$D_m \text{ de Gil será de: } D_d = \frac{|20-20|+|15-20|+|25-20|}{3} = \frac{0+5+5}{3} = \frac{10}{3} = 3,333\dots$$

$$D_m \text{ de Cleiton será de: } D_d = \frac{|30-20|+|25-20|+|5-20|}{3} = \frac{10+5+15}{3} = \frac{30}{3} = 10$$

Observamos que a média dos desvios de Gil é bem menor que a de Cleiton, isso significa que em média Gil está mais próximo da sua média de gols, há uma menor variação, enquanto isso não acontece tanto com Cleiton, cujo sua variabilidade é maior.

2.5.3 Variância

Para calcularmos a variância de um conjunto, que consiste em ser a média dos quadrados da diferença entre a média do conjunto e dos valores, podemos utilizar a fórmula e a representação:

$$S^2 = \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Em caso dos dados estarem em distribuição de frequência utilizaremos:

$$S^2 = \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 F_i}{n}$$

Exemplo 2.5.3. Consideremos a mesma situação do exemplo anterior, ou seja, as mesmas pessoas e as mesmas quantidades de gols, porém, vamos calcular a variância.

A variância de Gil será de:

$$S^2 = \frac{(20-20)^2+(15-20)^2+(25-20)^2}{3} = \frac{0^2+5^2+5^2}{3} = \frac{0+25+25}{3} = \frac{50}{3} = 16,666\dots$$

E a variância de Cleiton será de:

$$S^2 = \frac{(30-20)^2+(25-20)^2+(5-20)^2}{3} = \frac{10^2+5^2+15^2}{3} = \frac{100+25+225}{3} = \frac{350}{3} = 116,666\dots$$

2.5.4 Desvio Padrão

Em muitos casos, é preferencial o uso do desvio padrão e não o desvio médio devido ao modo de como se calcula, que, por sua vez, está diretamente relacionado com a variância. O desvio padrão sairá de imediato, pois, na variância, os valores

terão um resultado quadrático que pode ser resolvido com a extração da raiz quadrada. O resultado desse procedimento chamamos desvio padrão. Da mesma forma que encontramos o desvio padrão pela variância, encontraremos a variância pelo desvio padrão, a ordem depende do autor. Dizemos que o desvio padrão é:

$$\sqrt{\hat{\sigma}^2} = \hat{\sigma}$$

ou

$$\sqrt{S^2} = S$$

Exemplo 2.5.4. *Usando ainda o mesmo exemplo, podemos calcular o desvio padrão de Gil e Cleiton.*

O desvio padrão de Gil é:

$$\sigma = \sqrt{16,666\dots} \approx 4,08$$

Pare descobrir o desvio padrão de Cleiton, tem-se:

$$\sigma = \sqrt{116,666\dots} \approx 10,8$$

Podemos concluir que mesmo com a mesma média, os conjuntos não têm o mesmo comportamento, nos quais obedecem um desvio. Quanto menor o desvio, mais regular é este conjunto e quanto maior o desvio, menos regular será o conjunto, podendo assim entender, com mais variáveis, como um conjunto numérico funciona.

Capítulo 3

Comparação Estatística

Neste capítulo, faremos uma comparação estatística dos Campeonatos Brasileiros de Futebol da Série A e ilustraremos, desde que começou o formato de pontos corridos, com 20 clubes na disputa. O objetivo é fazer com que o leitor entenda a comparação, ao utilizar a estatística, e como podemos analisar esses resultados na prática, visto que essa abordagem pode ser utilizada não apenas no futebol, mas em outros esportes.

Apresentaremos as comparações mais comuns, a média de pontos para um time ser campeão brasileiro, a pontuação mágica para não ser rebaixado, a média de gols por time, rendimento dentro e fora de casa, etc. Além de como tudo isso influencia nas outras edições do campeonato, tomando como meta os valores encontrados. Antes de tudo, precisamos fazer uma breve introdução histórica de funcionamento da Série A do Campeonato Brasileiro de Futebol.

3.1 Série A do Campeonato Brasileiro de Futebol

O campeonato brasileiro de futebol, como conhecemos, começou a ser disputado em 1971 mas só começou a ser disputado no formato chamado "pontos corridos" em 2003, com 24 clubes. A presente dissertação terá como base, os campeonatos a partir de 2006, quando começou a ser disputado por 20 clubes, funcionando assim até hoje, por motivos de manter o padrão pela quantidade de jogos, fazendo com que assim os dados tenham um padrão de comportamento .

Usaremos dos dados das edições de 2006 até 2020, gerando, assim, um total de 15 edições para análise e comparação. O formato de pontos corridos, todos os 20

clubes se enfrentam duas vezes, cada uma no mando de campo de cada um dos clubes, totalizando 38 rodadas, 760 jogos e vários dados para serem coletados ao longo do torneio.

Um fato importante de ser exposto no trabalho é que cada vitória dentro do campeonato o time somará 3 pontos, cada empate somará 1 ponto e cada derrota não é adicionado nenhum ponto na competição. É importante também falar do “saldo de gols”, que é a diferença entre o número de gols marcados e o número de gols sofridos, que será negativo, caso um time sofra mais gols do que marque.

3.2 Comparações Estatísticas

Vamos começar a fazer a comparação da estatística dentro dos campeonatos, tomando como base as tabelas de classificações dos campeonatos brasileiros, após suas 38 rodadas, que ficarão em anexo, ao final do trabalho. Os valores serão todos baseados nas classificações dos campeonatos brasileiros da série A de 2006 até 2020 retiradas do site FlashScore e da CBF. Além disso, algumas contas serão desenvolvidas no corpo do trabalho e outras serão feitas com auxílio do Excel, mas todas foram verificadas com fórmulas do Excel. A ferramenta utilizada será apresentada mais à frente, no capítulo 4.

Uma aplicação similar a que faremos nesse trabalho é vista no filme Moneyball (O Homem que Mudou o Jogo) de 2011, estrelado pelo ator Brad Pitt, dirigido por Bennett Miller e com 6 indicações ao Oscar de 2012. Ele conta a história do gerente Billy Beane, do Oakland Athletics, time de Baseball dos Estados Unidos da América. Durante a temporada de 2002, perdendo seus 3 principais jogadores e com a 3ª menor folha salarial, o Billy foi em busca de jogadores utilizando um método estatístico, que avaliava a performance dos jogadores, tal como, o número de rebatidas e velocidade do jogador para chegar na base. Isso fazia com que ele criasse novos indicadores de medidas, buscando jogadores que apresentassem um bom rendimento nesses indicadores, fazendo assim com que o time contratasse jogadores mais baratos e com algum potencial.

3.2.1 Campeões Brasileiros

Uma das aplicações mais esperadas e desejadas é a média de pontuação que um clube precisa para ser campeão, com ela acompanham todas as variações que poderão acontecer e influenciar um clube a chegar nessa pontuação. Para que isso comece a fazer sentido, tomemos a tabela 3.2, criada apenas com os campeões brasileiros e suas respectivas pontuações.

Tabela 3.1: Pontuação dos campeões brasileiros

Pontuação	Time
90	Flamengo (2019)
81	Corinthians (2015)
80	Palmeiras (2018)
80	Palmeiras (2016)
80	Cruzeiro (2014)
78	São Paulo (2006)
77	São Paulo (2007)
77	Fluminense (2012)
76	Cruzeiro (2013)
75	São Paulo (2008)
72	Corinthians (2017)
71	Fluminense (2010)
71	Flamengo (2020)
71	Corinthians (2011)
67	Flamengo (2009)

Fonte: FlashScore

Percebamos que pela tabela têm valores variando de 67 pontos a 90 pontos, mesmo sabendo que “aparentemente” a maioria dos valores estão por volta de 70 a 80, para poder afirmar, precisamos calcular a média.

- $M = \frac{90+81+80+80+80+78+77+77+76+75+72+71+71+71+67}{15} = \frac{1146}{15} = 76,4$ pontos;
- A moda é $M_o = 71$ e 80 pontos, ocorrendo cada um em 3 ocasiões, sendo um conjunto bimodal;
- A mediana é $M_d = 77$ pontos, para isso, basta achar a posição da mediana com um conjunto de 15 elementos, $\frac{n+1}{2} = \frac{15+1}{2} = 8$ termos.

Ou seja, os 15 times campeões durante esses 15 anos, tiveram com uma pontuação mágica para o título de 76,5 pontos, porém observamos algumas situações com essa média e vamos discuti-la após calcular e avaliar a variação que ocorreu das pontuações do campeões, por isso vamos calcular o desvio padrão:

Figura 3.1: Cálculo das medidas por meio do excel

	A	B
1	MÉDIA	76,4
2	MODA	80
3	MEDIANA	77
4	DESVIO MÉDIO	4,24
5	DESVIO PADRÃO	5,64
6	VARIÂNCIA	31,83
7		
8	CLUBES	VALORES
9	FLAMENGO (2019)	90
10	CORINTHIANS (2015)	81
11	PALMEIRAS (2018)	80
12	PALMEIRAS (2016)	80
13	CRUZEIRO (2014)	80
14	SÃO PAULO (2006)	78
15	SÃO PAULO (2007)	77
16	FLUMINENSE (2012)	77
17	CRUZEIRO (2013)	76
18	SÃO PAULO (2008)	75
19	CORINTHIANS (2017)	72
20	FLUMINENSE (2010)	71
21	FLAMENGO (2020)	71
22	CORINTHIANS (2011)	71
23	FLAMENGO (2009)	67

Fonte: <https://www.flashscore.com.br/>

Consideremos que o desvio padrão vale 5,64 pontos, ou seja, dizemos que é um desvio baixo, significando que os valores estão em tornos da média por mais ou menos esse valor. Destacando que o cálculo foi feito com auxílio do Excel, que será um tema abordado apenas no capítulo 4, trazendo como podemos montar as tabelas que nos auxiliem em cálculos das medidas de tendência central e de dispersão dentro do objeto de estudos.

A pontuação que deverá ser feita por um time para garantir algum objetivo, observada com a aplicação no campeonato brasileiro, pode ser calculada apenas pela média, que para ser campeão foi de 76,4 pontos. No entanto, existem algumas variações como, por exemplo, entre os 15 campeões, 6 venceram com uma pontuação menor do que a média, sendo o pior deles o Flamengo de 2009, sendo essa média de 67 pontos.

Para ser campeão é que nunca na era dos pontos corridos com 20 clubes, um time alcançou mais do que 76,5 pontos e não foi campeão, o clube que mais chegou próximo foi o Santos (2º) e o Palmeiras (3º) de 2019, ambos com 74 pontos, ou seja, em 100% dos casos o número mágico funciona quando um time atinge a média, mas não quer dizer que um time com pontuação abaixo da média não seja campeão.

É válido ressaltar que para um time conseguir atingir uma pontuação de campeão terão vários fatores que implicarão naquele resultado, vamos expor mais alguns desses fatores, abordando por meio da estatística básica, estudada no ensino médio.

3.2.2 Clubes rebaixados para a Série B

O rebaixamento também se torna algo que interessa muitas pessoas, pois, no formato da Série A do campeonato brasileiro, os 4 últimos são rebaixados para a Série B do campeonato do ano seguinte e, como em toda disputa esportiva, sempre tem as equipes com menor investimento, que acabam lutando na parte debaixo da classificação, ou seja, os times que ficarão com as últimas colocações, isso não é diferente no futebol. Contudo, da mesma forma que funciona com o campeão, também existe um número mágico para o não rebaixamento. Alguns estudiosos da área indicam o número 46, visto que, nunca, na história da era pontos corridos com 20 clubes, um time com 46 pontos foi rebaixado. Porém nós vamos trabalhar com os fatos, ou seja, com o cálculo das suas medidas de tendências centrais e suas medidas de variações.

Construiremos uma Tabela 3.2 com o melhor time entre os 4 rebaixados, no caso os clubes que terminaram em 17º para conseguir fazer uma análise da pontuação para escapar do rebaixamento, visto que se o 17º terá a maior pontuação dos rebaixados. Lembrando que são rebaixados o 17º, 18º, 19º e 20º.

Tabela 3.2: Pontuação dos 17º colocados

Pontuação	Time
45	Coritiba (2009)
44	Portuguesa (2013)
44	Figueirense (2008)
44	Corinthians (2007)
43	Internacional (2016)
43	Coritiba (2017)
42	Sport (2018)
42	Avaí (2015)
42	Vitória (2010)
41	Sport (2012))
41	Athlético PR (2011)
41	Vasco (2020)
39	Ponte Preta (2006)
39	Vitória (2014)
36	Cruzeiro (2019)

Fonte: FlashScore

Como foi dito, nenhum dos clubes rebaixados chegaram aos 46 pontos, o mais próximo foi o Coritiba em 2009, que foi rebaixado na briga contra o Fluminense que ficou em 16º com os exatos 46 pontos. Mas essa pontuação a ser alcançada é suficiente, mas não necessária, pois em algumas edições tivemos times com menos de 46 pontos que não caíram. Para fazer essa análise com mais detalhes, traremos a tabela 3.3 com os 16º colocados, por ser o 1º time acima dos 4 rebaixados, ou seja, a menor pontuação dos times que permaneceram na Série A.

Tabela 3.3: Pontuação dos 16° colocados

Pontuação	Time
46	Fluminense (2009)
45	Goiás (2007)
45	Vitória (2016)
45	Flamengo (2013)
45	Portuguesa (2012)
44	Palmeiras (2006)
44	Naútico (2008)
43	Cruzeiro (2011)
43	Vitória (2017)
43	Figueirense (2015)
43	Vasco (2018)
42	Atlético GO (2010)
41	Fortaleza (2020)
40	Palmeiras (2014)
39	Ceará (2019)

Fonte: FlashScore

Observe a confirmação da frase que foi dita, que os 46 pontos eram suficientes para escapar do rebaixamento, porém não é necessário, pois dos 15 clubes que ficaram na 16° colocação apenas 1 deles atingiu os 46, todos os outros escaparam de ser rebaixado com uma pontuação menor.

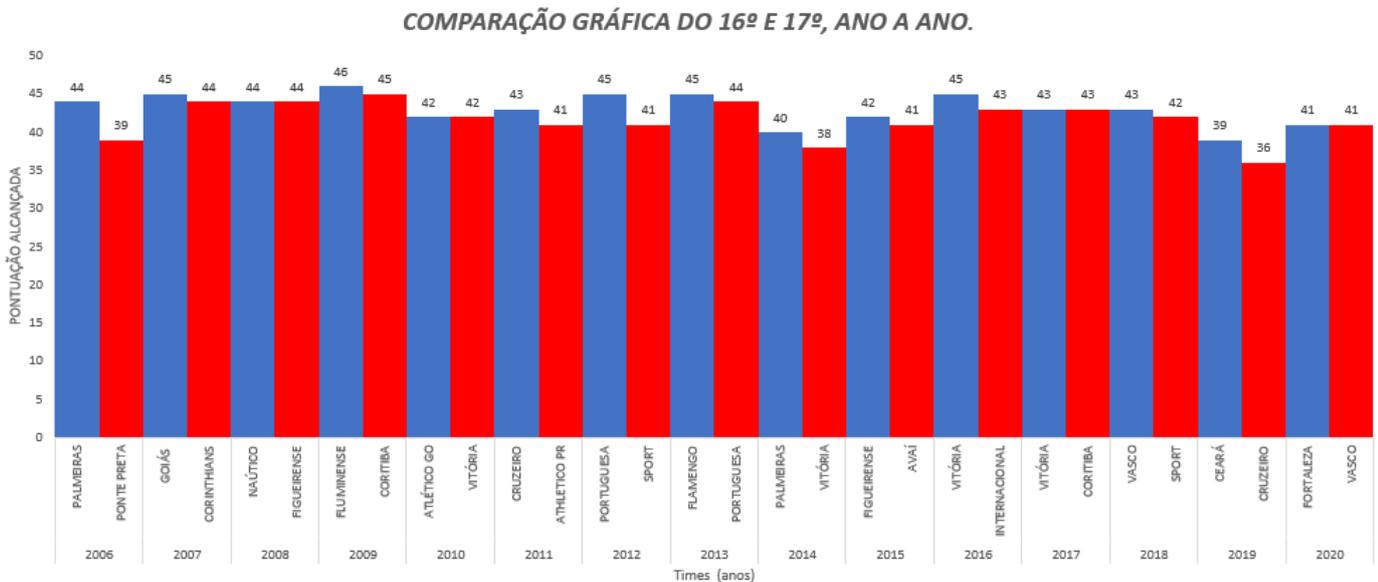
Antes de trazer um gráfico comparando as pontuações entre o 16° e o 17°, calcularemos suas respectivas médias.

$$M_{17} = \frac{45+44+44+44+43+43+42+42+42+41+41+41+39+39+36}{15} = 41,73 \text{ pontos para o } 17^\circ.$$

$$M_{16} = \frac{46+45+45+45+45+44+44+43+43+43+43+42+41+40+39}{15} = 43,20 \text{ pontos para o } 16^\circ.$$

Vejam que há uma diferença de 1,47 pontos, entre as médias dos último colocado não rebaixado e do primeiro rebaixado, ou seja, do 16° para o 17°, mas quando colocamos em prática, comparando a diferença de pontuação entre esses clubes que ficaram e os que caíram, como veremos na figura 3.2. Os clubes que ficaram em 16° estão em azul e os que ficaram em 17° em vermelho.

Figura 3.2: Comparação entre o 16º e o 17º colocado



Fonte: <https://www.flashscore.com.br/>

Com a comparação por meio de gráfico de barras conseguimos observar a diferença de pontuação dos clubes com mais detalhe. A maior diferença entre um permanente e um rebaixado é de 5 pontos, no caso do Palmeiras e Ponte Preta, no ano de 2006. Já a menor, foi de 0 pontos que ocorreu em 4 ocasiões, tendo seu rebaixado definido pelo número de vitórias ou saldo de gols, 2 temas que ainda iremos abordar. A média das diferenças entre os rebaixados e não rebaixados é de 1,47, como já dito antes. Concluimos, assim, que um time é sempre rebaixado por um detalhe, por um jogo devido a uma compilação de resultados durante todo o campeonato.

3.2.3 O grupo dos G4

O G4 é um termo utilizado para os 4 primeiros colocados do campeonato, que geralmente garantia vaga para a Libertadores (Campeonato a nível continental), mas com o passar dos tempos essas vagas foram crescendo e o G4 virou G7, porém apenas os 4 primeiros do campeonato brasileiro garantem vaga diretamente no torneio os outros 3 vão para a chamada Pré-Libertadores. O termo G4 é utilizado até hoje, até pelo fato de 4 caírem para a série B e 4 subirem, a mesma coisa acontecendo da B pra C e da C pra D, fazendo assim com que o G4 ainda seja algo para estudar o seu comportamento.

Para fazermos essa análise apresentaremos a Tabela 3.4 com os quartos colocados, servindo de pontuação mínima para conseguir ficar no G4.

Tabela 3.4: Pontuação dos 4º colocados

Pontuação	Time
69	Corinthians (2014)
66	São Paulo (2012)
66	Grêmio (2018)
66	São Paulo (2020)
65	Palmeiras (2008)
65	Grêmio (2019)
64	Santos (2006)
63	Grêmio (2010)
62	Grêmio (2017)
62	São Paulo (2015)
62	Cruzeiro (2009)
62	Atlético-MG (2016)
61	Botafogo (2013)
61	Fluminense (2007)
61	Flamengo (2011)

Fonte: FlashScore

Observamos as pontuações com mais clareza e concluímos que a pontuação que um time precisa para ficar em um grupo, é a menor delas. Calcularemos sua média, servindo como valor mínimo necessário:

$$M_{G4} = \frac{69+3*66+2*65+64+63+4*62+3*61}{15} = 63,67 \text{ pontos.}$$

O desvio padrão está em torno de 2,41 pontos, fazendo com que assim um time em torno de 64 pontos, arredondando, alcance a pontuação mínima para entrar nesse grupo. Podemos utilizar essa conta para um grupo maior do que os 4 primeiros, basta tomar uma tabela com a menor posição desse grupo, da mesma forma que foi aplicada com o G4, tal como o G7, basta pegarmos a pontuação dos 7º colocados.

Existe uma variação, com isso acaba ocorrendo alguns pontos fora da curva, ou seja, times que podem atingir a pontuação para pertencer ao grupo e não pertença, se tornando assim uma pontuação não 100% suficiente. O Fluminense de 2020, o

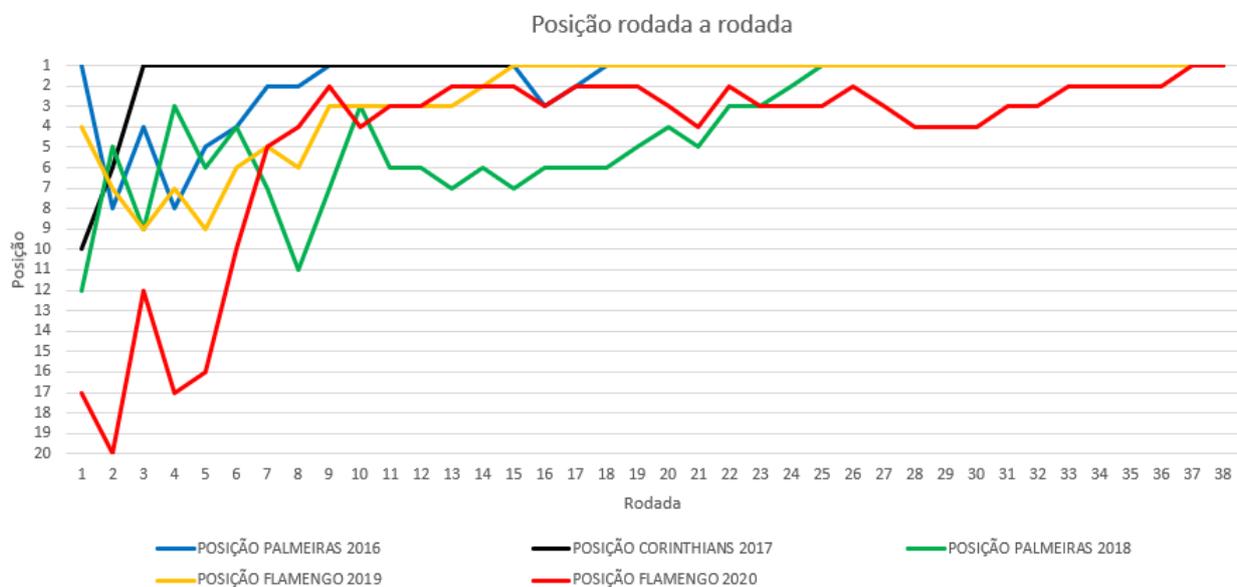
Athletico-PR de 2019 e o Flamengo de 2008 atingiram 64 pontos e ficaram apenas em 5º lugar. Com essa variação precisaremos considerar a pontuação mágica como sendo 65 pontos, para garantir o time entre os 4.

Passamos da análise dos campeões e dos grupos dos rebaixados que utilizam ideia de um valor máximo e de um valor mínimo, respectivamente, vamos abordar algumas outras situações que podem ser analisada e comparadas por meio estatística, buscando atingir os objetivos que possam ser traçados com base nos resultados obtidos.

3.2.4 Comparação entre campeões

Para iniciarmos esse tópico faremos a comparação entre os campeões, de forma a ver as semelhanças que existem nas campanhas dos times campeões, buscando descobrir qual o diferencial de cada time em relação ao outro e qual padrão é seguido ao longo dos anos. A figura 3.3 mostrará as posições dos últimos 5 campeões (2016-2020) rodada a rodada.

Figura 3.3: Posição dos campeões brasileiros de 2016-2020

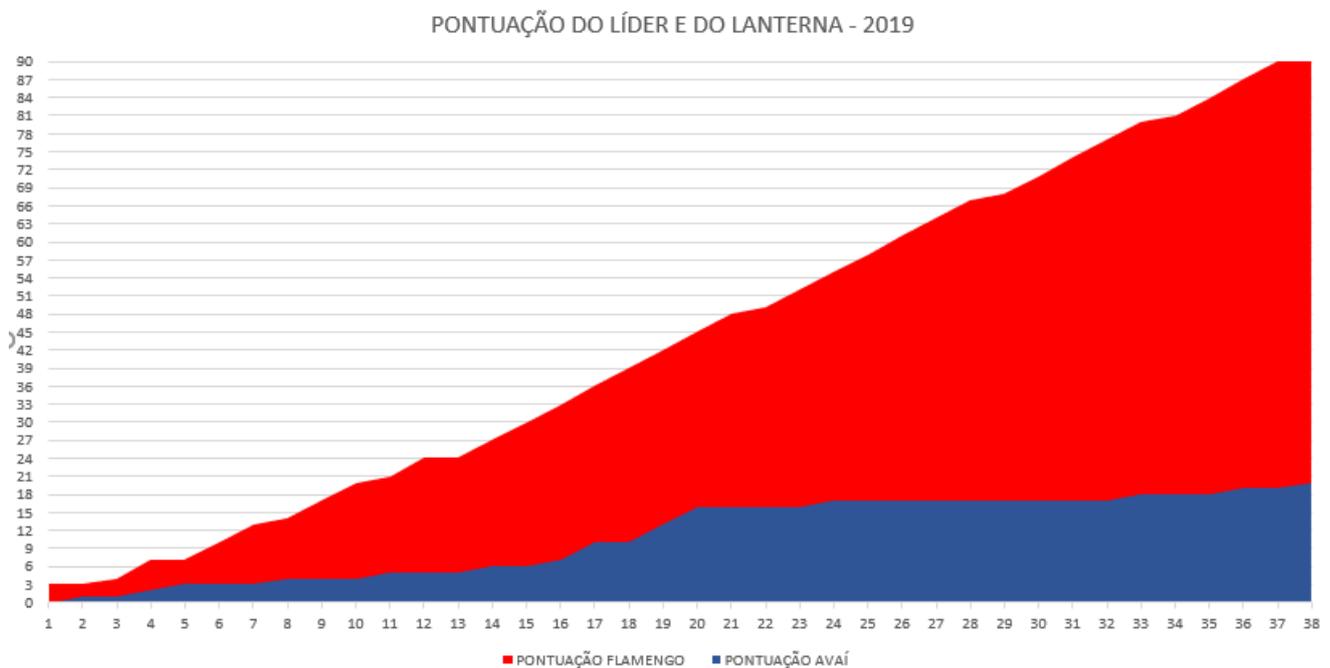


Fonte: Globo Esporte

Observamos nesse gráfico de linha o rendimento dos clubes rodada a rodada e posição a posição até chegar ao título. Verifica-se que, inicialmente, os clubes sempre buscam se manter-se entre os primeiros colocados, com exceção do Flamengo de

2020, que apenas conseguiu chegar entre os 4 primeiros na 8ª rodada. Outro fato que podemos observar é que todos os times, em algum momento, assumem a liderança, e mantêm-na ao longo da competição, com ressalva o Flamengo de 2020 que assumiu o 1º lugar na 37ª rodada. Esse é um gráfico muito bom para ver o rendimento dos clubes ao longo do campeonato, não somente entre campeões, mas entre quaisquer clubes. Outra forma de trazermos o rendimento de um clube é por meio do gráfico de áreas, no qual poderemos utilizar o avanço da pontuação de um time e sua pontuação acumulada durante o torneio.

Figura 3.4: Pontuação acumulada do Flamengo de 2019 e Avaí de 2019

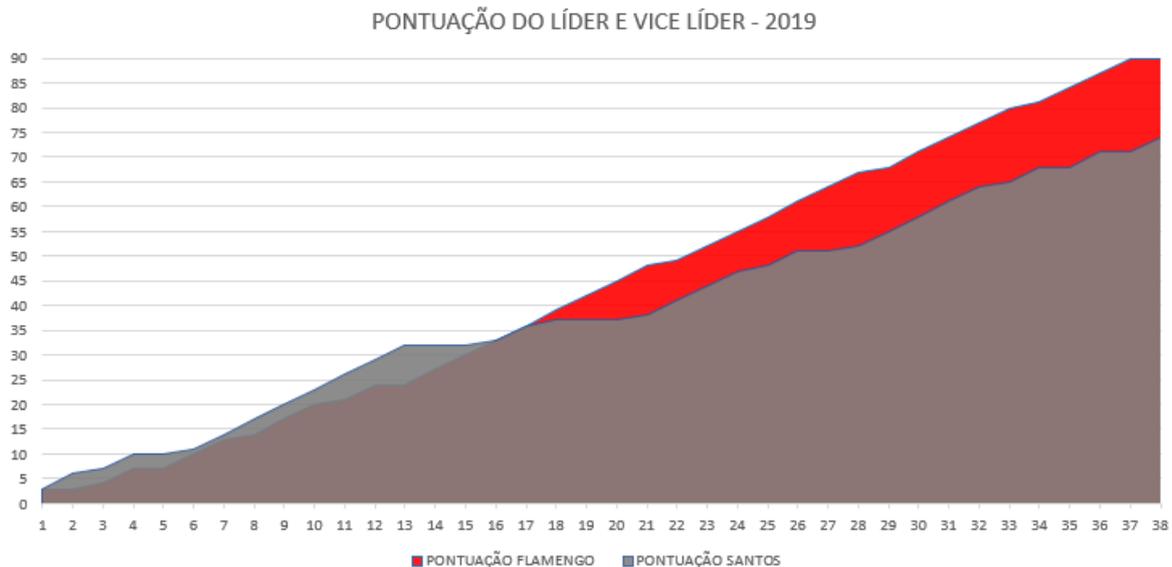


Fonte: SporTV

O gráfico acaba não se tornando tão bom para fazer comparações entre times, pois a área de um pode se sobressair a área do outro, caso esses times tenham desenvolvimentos parecidos. O gráfico apresenta o rendimento do Flamengo, campeão de 2019 e do Avaí, último colocado de 2019. Observemos que como os desenvolvimentos foram bem diferentes fica fácil vê-los em um único gráfico. É importante destacar que se o crescimento se mantiver constante significa que não somou pontos, ou seja, perdeu o jogo, se crescer apenas um ponto significa que empatou e se crescer três quer dizer que ganhou. Com isso conseguimos observar o desenvolvimento do clube ao longo do torneio, com o Flamengo tendo um crescimento ascendente durante boa

parte do campeonato e o Avaí uma evolução quase constante, significando muitos empates e derrotas ao longo do torneio.

Figura 3.5: Pontuação acumulada do Flamengo de 2019 e Santos de 2019



Fonte: SporTV

Do mesmo formato que a Figura 3.4 foi apresentada, a figura 3.5 também será. Entretanto, como já dito, utilizar times cujas pontuações são parecidas complicará mostrá-la com gráfico de áreas devido à sobreposição. No entanto, é importante observar a luta entre Flamengo e Santos, em que o clube Paulista ficou com pontuação à frente até a 16ª rodada, quando o próprio Flamengo ultrapassou e, a partir disso, apresentou uma maior vantagem em relação ao vice-campeão. Nesse contexto, procuramos utilizar a ferramenta que deixa a cor um pouco mais “transparente”, para que fique mais fácil diferenciar as pontuações, pois uma pergunta frequente não é *como* o time teve seu desenvolvimento ao longo do campeonato, mas *como* ele chegou a esse rendimento, bem como quantos gols marcou, quantos sofreu, somado a quantos jogos passou sem sofrer gols, entre outros fatores que pode levar um time ao fracasso. Veremos isso nas próximas seções

3.2.5 Número de gols marcados

Para iniciarmos, mostraremos o número de gols marcados por cada campeão, buscando expressar por meio da frequência, na tabela 3.5.

Tabela 3.5: Número de gols marcados

Gols marcados	Frequência
86	1
77	2
71	1
68	1
67	1
66	1
64	1
62	2
61	1
58	1
55	1
53	1
50	1

Fonte: FlashScore

No futebol, fazer gols é fundamental para conquistar vitórias. Quando mostramos o número de gols que um time campeão marca, com esse valor variando de 50 a 86 gols e uma média de 65,1 gols, fica vaga a ideia da relação entre gols marcados e o sucesso de um time. Isso ocorre devido ao fato de que nem sempre o campeão é o time que marca mais gols. Das 15 ocasiões analisadas, 8 delas obtiveram o time campeão como aquele que mais marcou gols, fazendo, assim, um aproveitamento em torno de 53,3%. Já nas outras 7, um time não campeão marcou mais gols do que o campeão, fato ocorrido em 2007, 2008, 2009, 2010, 2011, 2012 e 2017, o que pode ser verificado nas tabelas em anexo. Por este motivo, apresentaremos também as defesas dos campeões, na tabela 3.6.

Tabela 3.6: Número de gols sofridos pelos campeões

Gols sofridos	Frequência
48	1
44	1
38	1
37	2
36	3
33	1
32	2
31	1
30	1
26	1
19	1

Fonte: FlashScore

Com isso, podemos observar que a média de gol sofrida por um time campeão está em torno de 34,3 gols. No entanto, como também acontece com os gols marcados, acontece com os gols sofridos, em que nem sempre o time campeão foi o que sofreu menos gols, ou seja, não se pode concluir na implicação de que a melhor defesa será do campeão. Em 9 ocasiões o time campeão foi o que sofreu menos gols, com um aproveitamento de 60%, e, nas outras 6, outros times diferentes do campeão. Número, este, bem próximo dos melhores ataques, ocorrido em 2008, 2009, 2013, 2014, 2019 e 2020. Com isso, podemos concluir que não é garantia ter um bom ataque ou uma boa defesa para ser campeão. Para isso, falaremos do chamado “saldo de gols”, que é a diferença entre o número de gols marcados e o número de gols sofridos, expressa na Tabela 3.7.

Tabela 3.7: Saldo de gols dos campeões

Saldo de gols	Frequência
49	1
40	2
38	1
36	1
34	1
30	2
29	1
28	1
26	1
20	2
17	1
14	1

Fonte: FlashScore

Quando observamos o saldo de gols que tem uma média de 30,1, conseguimos ser mais precisos para mensurar do que um time precisa para possivelmente alcançar o título, ao buscar equilíbrio entre marcar mais gols e sofrer menos gols. Na tabela vemos que os saldos dos 15 clubes campeões, 13 tiveram os maiores saldos, ocorrendo duas exceções, 2009 e 2020. Isso resultou em um aproveitamento de 86,6% aproximadamente.

Uma outra indicação do que também pode levar um time ao título é passar jogos sem sofrer gols, o que é chamado no futebol de “Clean sheets”, termo em inglês. Nesse contexto, colhemos os dados dos campeões das 15 edições da Série A, dos jogos em que não sofreram gols, e observamos isso na tabela 3.8.

Um dado importante é que toda a coleta foi feita pelo site da CBF e da UOL, como também foi observada cada partida de cada clube no ano do seu título.

Tabela 3.8: Jogos sem sofrer gols dos campeões (Clean sheets)

Campeão	Ano do campeão	Clean sheets
Flamengo	2020	9
Flamengo	2019	14
Palmeiras	2018	18
Corinthians	2017	18
Palmeiras	2016	15
Corinthians	2015	17
Cruzeiro	2014	11
Cruzeiro	2013	15
Fluminense	2012	14
Corinthians	2011	14
Fluminense	2010	12
Flamengo	2009	16
São Paulo	2008	10
São Paulo	2007	22
São Paulo	2006	15

Fonte: Uol

Nessa análise observou-se que com o número de jogos sem sofrer gol, nas 38 partidas, temos que o clube campeão possui, em média, 14,67 dos seus jogos sem sofrer gols, com um desvio padrão de 3,26 jogos. Diante disso, podemos ver que o pior desempenho nesse quesito foi do time Flamengo, no campeonato de 2020. Dos 38 jogos, o time possuía apenas 9 jogos sem sofrer gols, sendo um valor de $\frac{9}{38} = 0,2368 = 23,68\%$. Já o melhor aproveitamento foi do time São Paulo, de 2007, que não sofreu gols em 22 dos 38 jogos, trazendo um valor de $\frac{22}{38} = 0,5789 = 57,89\%$ sem sofrer um gol. A partir disso, podemos dizer que esse desempenho de gols marcados e não sofridos, ajuda um clube, na maioria das vezes, a ser campeão.

3.2.6 Aproveitamento dos clubes

Para falar do aproveitamento dos clubes e mostrar que existe um padrão entre o campeões, introduziremos a ideia que foi utilizada para a construção das tabelas. Como sabemos, no campeonato há um total de 38 rodadas, sendo jogados 19 em mando de campo do time e 19 no mando de campo do time rival. Nesse contexto,

o aproveitamento é feito pela razão entre os pontos conquistados e os pontos que disputou, multiplicado por 100, para, então, gerar a porcentagem.

$$A = \frac{P_c}{P_d} * 100$$

Onde A é o aproveitamento, P_c são os pontos conquistados e P_d são os pontos disputados.

Fizemos uma análise de jogo a jogo, dos campeões, nos anos dos seus títulos, segundo o site da CBF [3] para os campeonatos de 2012 até 2020 e o site da UOL [24] para os campeonatos de 2006 até 2011. Trouxemos o número de vitórias, empates e derrotas, dentro e fora de casa, separadamente. Lembrando que uma vitória vale 3 pontos, um empate vale 1 ponto e a derrota vale 0 pontos, com isso, dos 19 jogos em casa e 19 fora, há em disputa $3 * 19 = 57$ em casa e outros 57 fora, totalizando 114 pontos em disputa, ou seja, o número total de pontos que cada time poderá alcançar. Observemos a tabela 3.9.

Tabela 3.9: Aproveitamento dos campeões jogando em casa

Campeão	Ano do campeão	Vitória	Empate	Derrota	Aproveitamento
Flamengo	2020	11	4	4	64,91%
Flamengo	2019	17	2	0	92,98%
Palmeiras	2018	16	2	1	87,72%
Corinthians	2017	12	2	5	66,67%
Palmeiras	2016	14	4	1	80,70%
Corinthians	2015	16	2	1	87,72%
Cruzeiro	2014	15	2	2	82,46%
Cruzeiro	2013	14	3	2	78,95%
Fluminense	2012	11	5	3	66,67%
Corinthians	2011	13	2	4	71,93%
Fluminense	2010	12	5	2	71,93%
Flamengo	2009	12	5	2	71,93%
São Paulo	2008	14	4	1	80,70%
São Paulo	2007	13	3	3	73,68%
São Paulo	2006	14	4	1	80,70%

Fonte: CBF e Uol

Com a Tabela 3.9, fica visível o rendimento dos times campeões quando jogam

diante de sua torcida. Isso mostra que eles sempre vencem mais da metade de seus jogos quando realizados no seu mando de campo (ou seja, “em casa”), possuem, em média, um aproveitamento de 77,31% (calculado com auxílio do Excel), mas possuem o pior rendimento em torno de 66,67%, ou seja, os times sempre conquistam mais do que $\frac{2}{3}$ dos seus pontos. Isso representa, pelo menos, 38 pontos dos 54 disputados. Pode-se, portanto, concluir que um bom rendimento jogando seu mando de campo é crucial na disputa pelo título. Isso demonstra um desvio padrão em torno de 8,19%, mas, ainda assim, mostra uma complementação e padronização dos rendimentos obtidos.

Para fazermos uma complementação dos rendimentos, mostraremos na tabela 3.10 o número de vitórias, empates, derrotas e rendimentos dos times ao jogarem no campo adversário. É válido lembrar que também são 19 jogos fora, com 54 pontos em disputa, cujo rendimento também é da pontuação conquistada em relação à pontuação total em disputa..

Tabela 3.10: Aproveitamento dos campeões jogando fora de casa

Campeão	Ano do campeão	Vitória	Empate	Derrota	Aproveitamento
Flamengo	2020	10	4	5	59,65%
Flamengo	2019	11	4	4	64,91%
Palmeiras	2018	7	9	3	52,63%
Corinthians	2017	9	4	6	54,39%
Palmeiras	2016	10	4	5	59,65%
Corinthians	2015	8	7	4	54,39%
Cruzeiro	2014	9	6	4	57,89%
Cruzeiro	2013	9	4	6	54,39%
Fluminense	2012	11	6	2	68,42%
Corinthians	2011	8	6	5	52,63%
Fluminense	2010	8	6	5	52,63%
Flamengo	2009	7	5	7	45,61%
São Paulo	2008	7	8	4	50,88%
São Paulo	2007	10	5	4	61,40%
São Paulo	2006	8	8	3	56,14%

Fonte: CBF e Uol

Veremos, então, que o rendimento no campo adversário diminui em comparação ao rendimento em seu mando de campo, tendo, assim, em média um rendimento

de 56,37%. Nesse contexto, apenas um dos campeões ficou abaixo dos 50% de rendimento jogando fora, que foi o Flamengo, em 2009, e, apenas uma vez, o rendimento do campeão foi melhor no campo rival, sendo, ele, o Fluminense de 2012. Conseguimos, então, perceber que se deve manter bons rendimentos, com o seu campo ou em campo adversário, ao se tentar sempre aproveitar seus jogos local. O desvio padrão está em torno de 5,58%, um pouco menor do que o desvio dos rendimentos *em casa*, mostrando mais ainda a constância dos jogos no campo do adversário.

De maneira geral, podemos mostrar o rendimento dentro de todo o campeonato ao juntar as duas tabelas apresentadas, fazendo, assim, a média dos rendimentos em casa e fora, por meio da média aritmética ou média geométrica, isso possibilita obter essa média geral na Tabela 3.11.

Tabela 3.11: Aproveitamento dos campeões

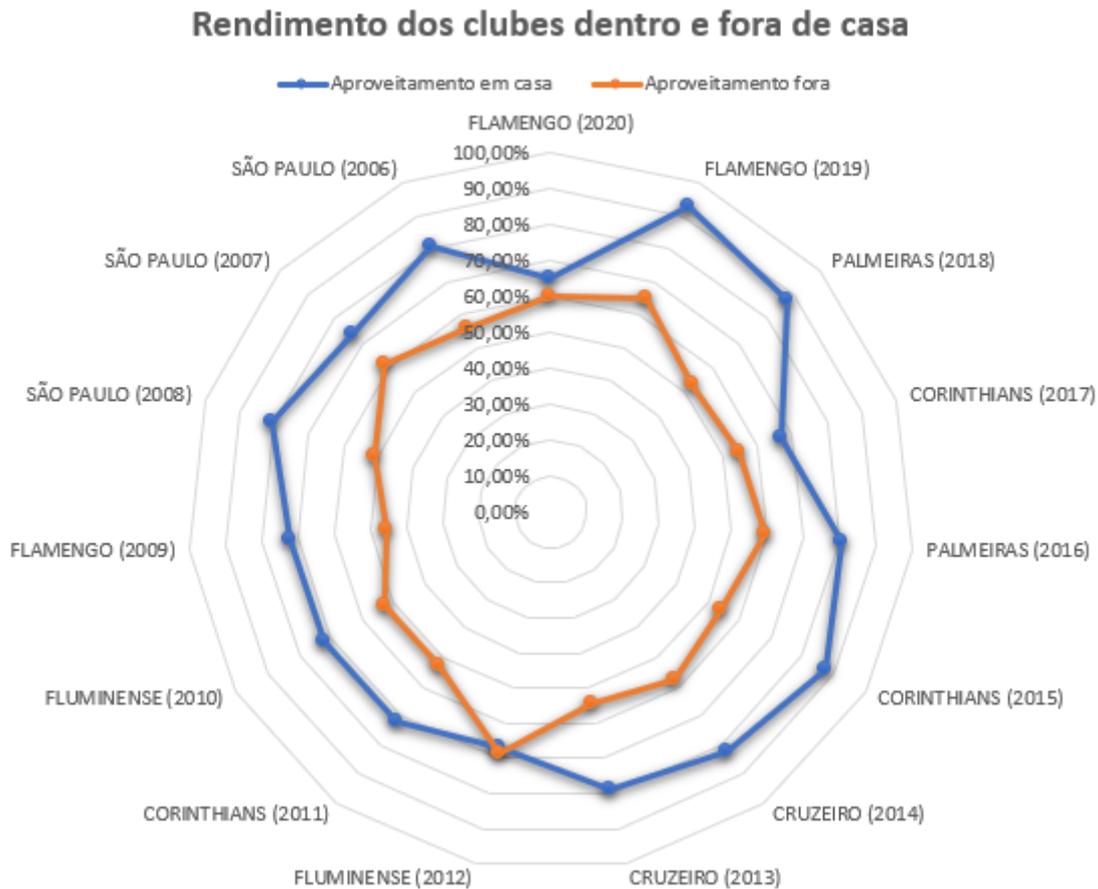
Campeão	Ano do campeão	Aproveitamento
Flamengo	2020	62,28%
Flamengo	2019	78,95%
Palmeiras	2018	70,18%
Corinthians	2017	60,53%
Palmeiras	2016	70,18%
Corinthians	2015	71,05%
Cruzeiro	2014	70,18%
Cruzeiro	2013	66,67%
Fluminense	2012	67,54%
Corinthians	2011	62,28%
Fluminense	2010	62,28%
Flamengo	2009	58,77%
São Paulo	2008	65,79%
São Paulo	2007	67,54%
São Paulo	2006	68,42%

Fonte: CBF e Uol

É possível avaliar que os rendimentos são todos em volta de 66,84%, sendo essa a média do campeão. Aqui, o desvio padrão é o menor das 3 tabelas, sendo ele por volta de 4,96%. Isso mostra que, para ser campeão, os rendimentos devem ser bem parecidos, desviando um valor em torno de 5%, que é equivalente a, aproximadamente,

5 pontos. Para mostrar, então, a diferença entre essas 3 tabelas, traremos um gráfico em radar na figura 4.2, para comparar os 2 rendimentos, ano a ano, time a time.

Figura 3.6: Rendimento dos clubes dentro e fora de casa



Fonte: CBF e UOL

Notemos, portanto, que o padrão nos rendimentos dos campeões, geralmente, está com uma distância média de 20% de diferença entre o rendimento com mando de campo e no campo adversário, algo que foi confirmado com os cálculos das médias. Observamos, também, a ocasião em que o rendimento no campo adversário foi melhor que o rendimento em casa, no ano de 2012. Podemos chamá-lo, no entanto, de ponto fora da curva, pois aconteceu apenas uma vez.

Com todas essas análises, fica visivelmente claro o rendimento que um time precisa obter. Com isso, sugerimos algumas atividades que podem ser aplicadas pelo professor aos alunos em sala de aula. Essas atividades podem ser modificadas de acordo com a necessidade e preferência do professor. Além disso, a comparação estatística pode ser aplicada em outros esportes ou atividades esportivas.

3.3 Sugestão de atividade

A seguintes sugestão tem como objetivo o desenvolvimento de atividades por parte dos alunos dentro da disciplina de Estatística.

- **Tema:** fazendo um resumo das principais comparações que podem ser feitas em algum campeonato ou esporte escolhido pelo aluno, o professor pode apenas direcionar qual tema o aluno deve escolher e quais informações poderia colher. Sugestão: pontuação dos campeões, número de vitória dos campeões, número de pontos ou gols marcados pelos campeões, etc.
- **Campeões:** os alunos podem buscar um campeonato de algum esporte e selecionar os campeões da quantidade de anos que for sugerido pelo professor. Quanto maior o número, melhor a análise estatística. Sugestão: solicitar ao menos 10 edições de um campeonato, para que o aluno colete os dados.
- **Tabelas no Excel:** deve ser solicitado aos alunos a criação de tabelas com as informações coletadas dentro do Excel. A vantagem que mais a frente os alunos podem usar tabela como fonte para criação de fórmulas para alguns cálculos.
- **Número de vitórias ou pontuação:** podem ser colhidos os dados acerca do número de vitórias ou pontuação, dependendo de qual esporte for escolhido pelo aluno. Com isso, será possível calcular a média, moda e mediana desses dados.
- **Gráficos:** com as informações colhidas e colocadas em uma tabela, os alunos deverão elaborar gráficos, comparando cada edição ou o desenvolvimento de um time.
- **Cálculo das medidas de tendência central:** devem ser feito os cálculos a partir dos dados colhidos pelos alunos, usando o Excel como ferramenta de auxílio para o cálculo das médias, variância e desvio padrão.

- **Conclusão:** deverá ser feita uma conclusão, com base nas informações, do que pode ser visto ou esperado durante outros campeonatos, visando o que um time poderá utilizar como base para traçar metas a serem alcançadas, baseadas naquele desempenho padrão visto pelas médias, desvios, gráficos, tabelas, etc.
- **Referência:** o referencial deve ser lembrado, no qual se informa de onde foram retirados os dados. Em qualquer informação, deverá ser colocada a fonte ao final.
- **Entrega:** a atividade pode ser entregue em documento pelos alunos, e, se for preferência do professor, poderá ser solicitada uma apresentação curta sobre os dados colhidos, o que foi construído e o que foi concluído.

Capítulo 4

Utilização do Excel

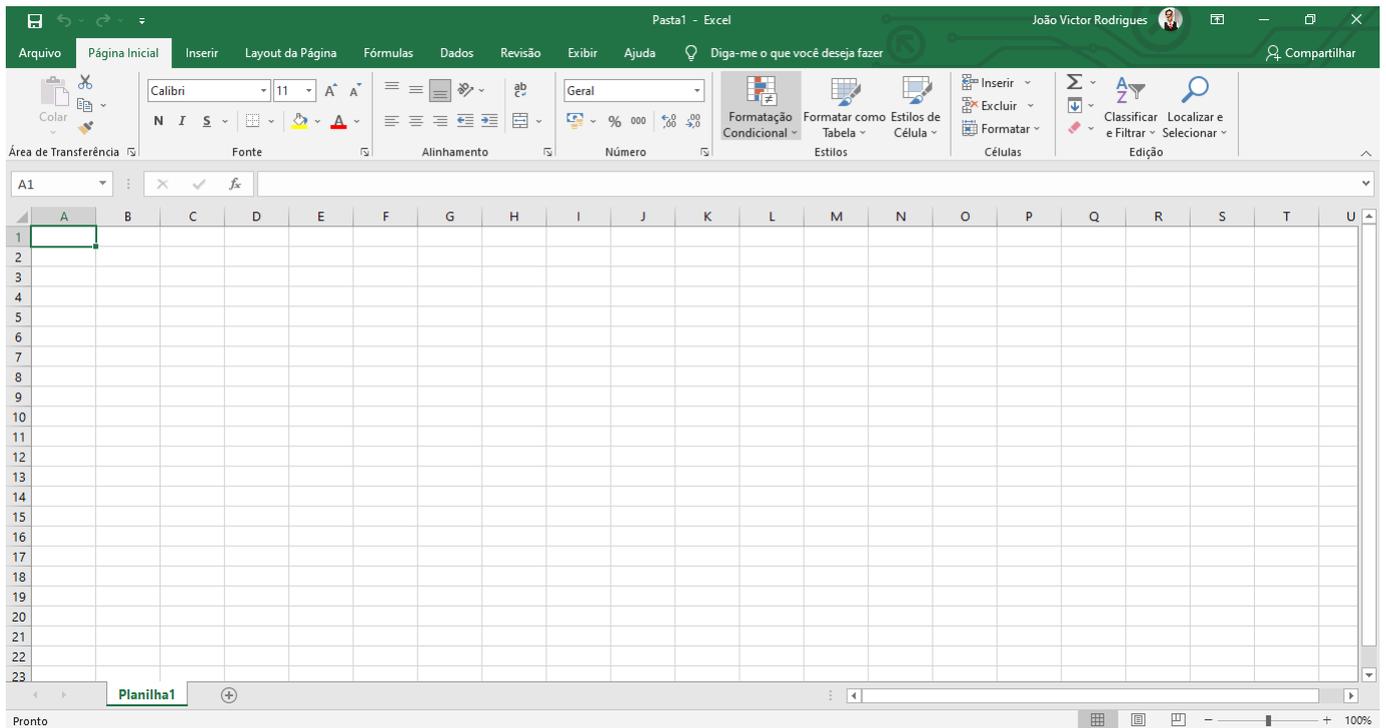
Neste capítulo, falaremos sobre a ferramenta presente em vários celulares e computadores: o Excel - aplicativo do pacote Office da Microsoft. Faremos uma introdução ao uso do aplicativo e mostraremos quais utilidades presentes nele poderão nos ajudar em uma comparação estatística, com a criação de tabelas, gráficos e, inclusive, cálculos de medidas de tendência central e medidas de dispersão. Segundo Levine ([14], p.28), dentre os programas de planilhas de cálculo eletrônicos o Microsoft se destaca pela acessibilidade e pela participação dominante.

Para Fonseca ([8], p.3), o Excel é um software de planilha eletrônica ou de cálculo, desde cálculos simples aos mais avançados. Sua importância, como ferramenta dentro da sala de aula, se dá para o desenvolvimento do raciocínio do aluno, para que ele entenda a suas utilidades no contexto proposto pelo professor, o que, para SILVA ([21], p.3), poderá aguçar o desejo desse aluno pelos assuntos ensinados, aliando uma metodologia aplicada pelo professor com um software que muitos já conhecem, ou, aqueles que não conhecem, precisarão desenvolver a habilidade de utilizar uma planilha para alguma situação do dia a dia, seja pessoal ou profissional.

4.1 O Excel

Apresentaremos algumas ferramentas do Microsoft Excel, a partir da sua tela inicial, informaremos suas funcionalidades básicas para que em sequência possamos detalhar.

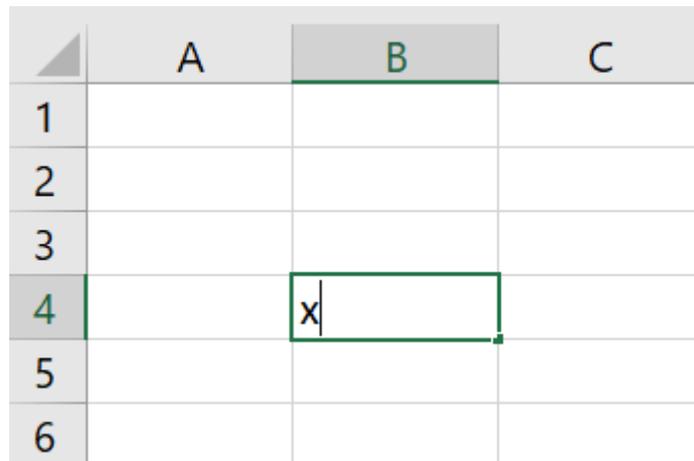
Figura 4.1: Aplicativo do Excel no computador



Ao analisarmos a página inicial do Excel, na figura 4.1, destacamos alguns pontos importantes que usaremos em nosso trabalho:

- **Células e barra de fórmulas:** Os retângulos, nos quais podemos inserir os números e nome, são as células, que estão ligadas diretamente à barra de fórmulas, cuja função é inserir os comandos. As células são identificadas por um número e uma letra. O número indicará qual linha está a célula e a letra indicará a coluna.

Figura 4.2: Célula no Excel



The image shows a portion of an Excel spreadsheet. The columns are labeled A, B, and C, and the rows are labeled 1 through 6. The cell at the intersection of column B and row 4 (cell B4) is selected, indicated by a green border. The letter 'x' is entered into this cell. The cell at the intersection of column A and row 4 (cell A4) is also highlighted with a grey background.

	A	B	C
1			
2			
3			
4		x	
5			
6			

Observemos que na figura 4.2 o x está na célula B4. Para inserir algo, basta clicar em cima com o mouse e após apertar na tecla enter.

- **Página inicial:** localizada na barra superior, ela traz funções que se aplicam nas letras, como as fontes, tamanhos, posições, formatos, estilo dos números e formatação das células.
- **Inserir:** é a parte do excel em que utilizamos ferramentas na qual nos permite inserir imagens, tabelas e gráficos, podendo ser considerada a ferramenta parte mais importante para o nosso trabalho.
- **Layout da página:** é a função em que podemos formatar toda a página, margens da folha, cores e qualquer outra edição que envolva o layout.
- **Fórmulas:** aba que apresenta as fórmulas mais usadas, entre elas as de Matemática financeira, trigonometria, data/hora, etc. É importante devido a todas as funções que utilizamos em nosso trabalho, com uso dos cálculos faremos análise de uma situação.

Feita essa apresentação básica sobre algumas ferramentas fundamentais no Excel, traremos os comandos para as criações que foram feitas em nosso trabalho e de como podemos utilizá-lo ao nosso favor.

4.2 Tabelas

A tabela é uma das principais ferramentas do Excel, usada para a organização dos dados brutos.

Figura 4.3: Rendimento dos últimos 5 campeões jogando em casa

	A	B	C	D	E
1	TIME (ANO)	CASA			
2		VITÓRIA	EMPATE	DERROTA	APROVEITAMENTO
3	FLAMENGO (2020)	11	4	4	64,91%
4	FLAMENGO (2019)	17	2	0	92,98%
5	PALMEIRAS (2018)	16	2	1	87,72%
6	CORINTHIANS (2017)	12	2	5	66,67%
7	PALMEIRAS (2016)	14	4	1	80,70%

Fonte: UOL e CBF

A Tabela na figura 4.3 está organizada por ordem dos anos, mas também poderia ser organizada em ordem de aproveitamento, do melhor para o pior.

A porcentagem é o valor que representa quantas partes de 100 o número tem. O primeiro passo é saber qual o total, ou seja, o time joga 19 vezes em casa e tem em disputa $3 * 19 = 57$ pontos, pois cada jogoo time poderá ganhar 3 pontos, em caso de vitória. Um modo de calcular, exemplificado com o Flamengo de 2020, é:

$$\frac{37}{57} * 100 = 64,91\%$$

Sabendo que o Flamengo ganhou 11 jogos, somando $3 * 11 = 33$ pontos e 4 empates, somando 37 pontos conquistados. Mas isso pode ser feito no Excel de maneira mais prática usando o código:

$$= [(B3 * 3 + C3 * 1)/57] * 100$$

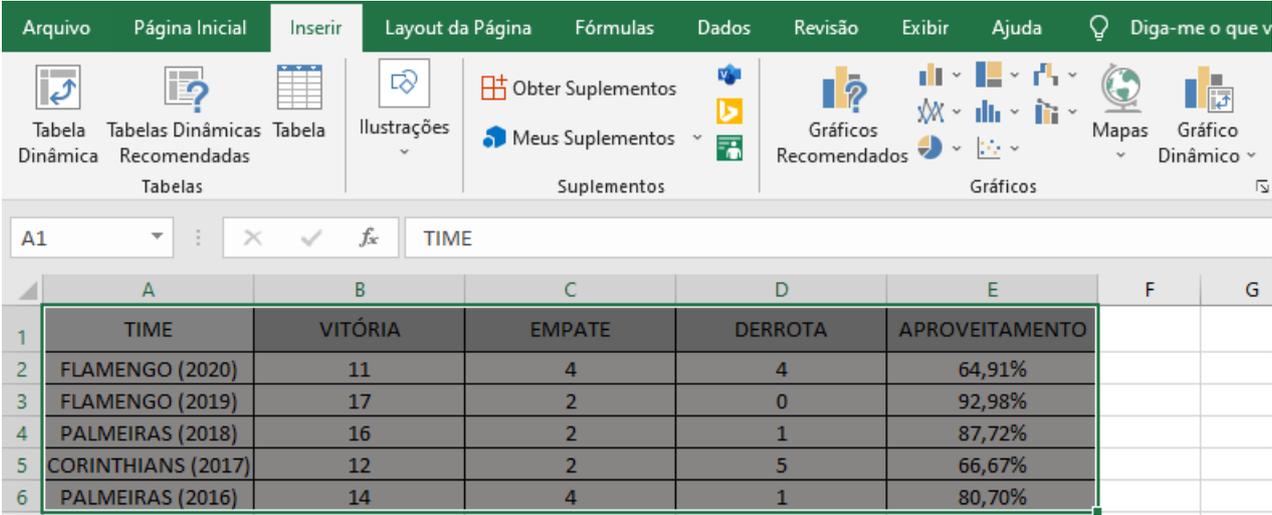
Sendo B3 a célula que tem o número de vitórias e C3 a célula que tem o número de empates, por isso estão multiplicados (*), respectivamente, por 3 e 1. O sinal de

“=” sempre deverá ser colocado antes de inserir alguma fórmula. Para a inserção das fórmulas nas linhas de baixo devemos clicar na célula que foi inserida, no caso a E3, e arrastar até onde desejar que ela também seja aplicada, no caso na E7.

4.3 Gráficos

Para uma outra representação dos dados tabelados usaremos os gráficos, que no trabalho aplicamos os gráficos em barras, de setores, de segmento, em área e em radar. O método para construção dos gráficos é basicamente o mesmo.

Figura 4.4: Inserir gráfico



	A	B	C	D	E	F	G
1	TIME	VITÓRIA	EMPATE	DERROTA	APROVEITAMENTO		
2	FLAMENGO (2020)	11	4	4	64,91%		
3	FLAMENGO (2019)	17	2	0	92,98%		
4	PALMEIRAS (2018)	16	2	1	87,72%		
5	CORINTHIANS (2017)	12	2	5	66,67%		
6	PALMEIRAS (2016)	14	4	1	80,70%		

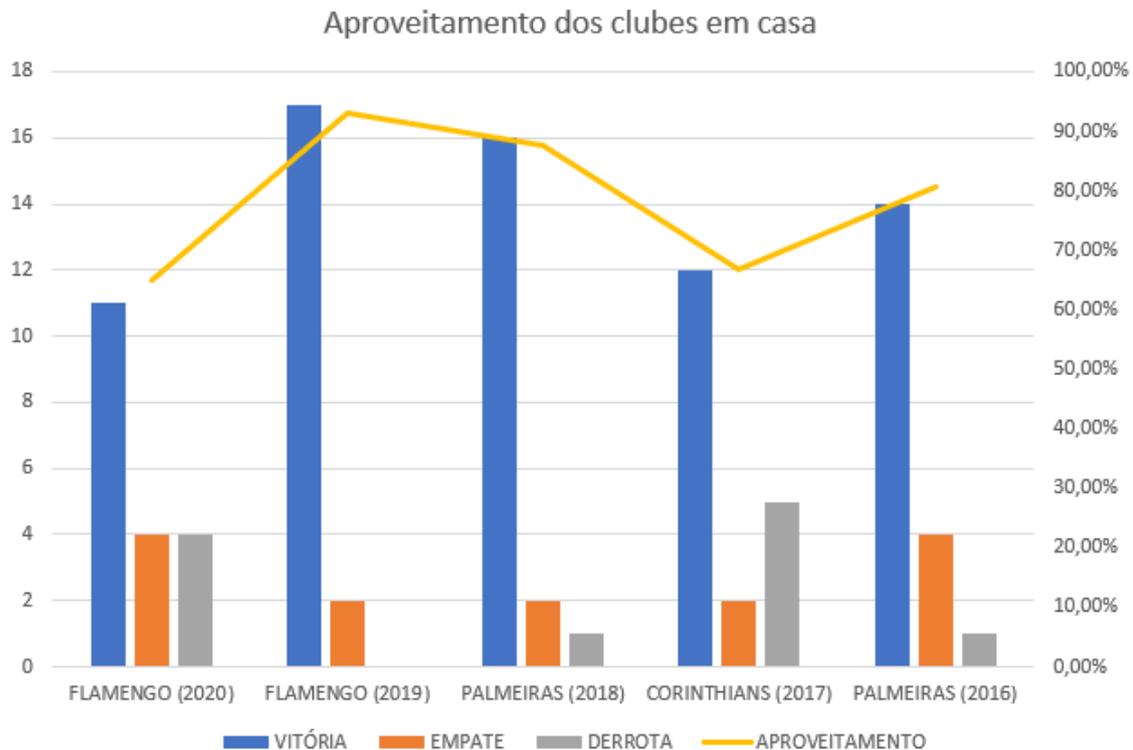
No processo, basta selecionarmos as informações na tabela, como feito na figura 4.4 e na aba de inserir clicarmos na opção de adicionar os gráficos. Existe a opção ir em gráficos recomendados ou direto no gráfico do seu interesse.

Figura 4.5: Gráficos recomendados



Teremos nessa parte os gráficos que ficam mais adaptados ao tipo de tabela e das informações que tivermos utilizado. Selecionaremos o gráfico que é uma mistura de gráfico em colunas, com as informações das vitórias, junto do gráfico de segmentos, com as informações sobre o aproveitamento dos clubes, como visto na figura 4.5. Em sequência basta clicar em confirmar.

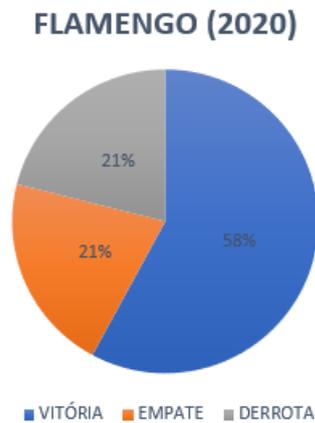
Figura 4.6: Aproveitamento dos clubes jogando em casa



O gráfico pronto aparece na Figura 4.6. A criação de gráficos é uma ferramenta totalmente fácil de se utilizar, com objetivo de nos mostrar uma outra visualização do padrão dos clubes, tanto no número de vitórias, empates e derrotas, além do aproveitamento. Existe a opção de editar cores, nomes e grandezas, basta clicarmos com o botão direito em cima de onde deseja editar terá acesso as configurações.

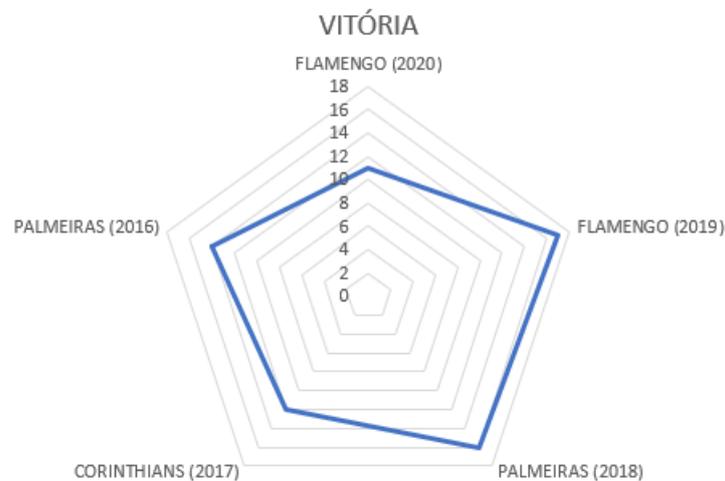
Há uma outra opção, a de mostrar apenas o número de vitórias, empates e derrotas de cada clube individualmente, usando, portanto, o gráfico de setores, ou gráfico de pizza. Para isso, basta selecionarmos os dados de um único clube e seguir o mesmo passo feito no gráfico anterior.

Figura 4.7: Aproveitamento do Flamengo, no ano de 2020, jogando em casa



Para inserirmos um gráfico de radar, podemos selecionar apenas o número de vitória dos clubes, que serão representados em um radar, comparando o número de vitória dos mesmos, como na Figura 4.8. Lembrando que isso também poderia ser representado um gráfico de barras, apenas com uma outra visualização.

Figura 4.8: Comparação do número de vitória dos últimos 5 campeões



Colocamos esses 3 gráficos por se encaixarem dentro do contexto que criamos entre os 5 últimos campeões, mas existem outros que utilizamos no trabalho, porém, a forma de inserir é análoga, selecionando as células com as informações e, em sequência, o gráfico mais viável para as variáveis que foram utilizadas, podendo ser formatada ao final.

4.4 Medidas de tendência central

Para os cálculos de medidas de tendência central é suficiente inserirmos a fórmula, a partir dos valores da tabela. Exemplificando com a tabela criada pelo Excel, aplicada na pontuação dos campeões brasileiros, exposta pela Figura 4.3.

Figura 4.9: Cálculo da média

	A	B
1	MÉDIA	
2	MODA	
3	MEDIANA	
4	DESVIO MÉDIO	
5	DESVIO PADRÃO	
6	VARIÂNCIA	
7		
8	CLUBES	VALORES
9	FLAMENGO (2019)	90
10	CORINTHIANS (2015)	81
11	PALMEIRAS (2018)	80
12	PALMEIRAS (2016)	80
13	CRUZEIRO (2014)	80
14	SÃO PAULO (2006)	78
15	SÃO PAULO (2007)	77
16	FLUMINENSE (2012)	77
17	CRUZEIRO (2013)	76
18	SÃO PAULO (2008)	75
19	CORINTHIANS (2017)	72
20	FLUMINENSE (2010)	71
21	FLAMENGO (2020)	71
22	CORINTHIANS (2011)	71
23	FLAMENGO (2009)	67

Para calcular a média na tabela, clicamos na célula B1 e inserimos na expressão:

$$= MÉDIA(B9 : B23)$$

Ao selecionarmos as células que têm os valores, obtemos o resultado da média imediatamente. Para selecionarmos os valores, basta o uso do mouse, clicando na B9 e arrastando até os últimos valores, apertando confirmando com um enter. A fórmula também poderá ser inserida apenas com o uso do teclado, digitando “B9:B23”.

Os cálculos da moda e mediana são feitos de maneira análoga, inserindo nas células B2 e B3, respectivamente, as expressões:

$$= \text{MODO}(B9 : B23)$$

$$= \text{MED}(B9 : B23)$$

Observemos que a moda é expressa por 2 valores, 80 e 71, mas o Excel vai trazer apenas um dos valores dessa moda, que seria o maior deles. Vejamos a inserção da fórmula na Figura 4.10.

Figura 4.10: Cálculo da mediana

	A	B
1	MÉDIA	76,4
2	MODA	80
3	MEDIANA	=MED(B9:B23)
4	DESVIO MÉDIO	
5	DESVIO PADRÃO	
6	VARIÂNCIA	

4.5 Medidas de dispersão

Para calcularmos as medidas de dispersão usaremos uma fórmula análoga, não muito diferente da usada nas medidas de tendência central. Utilizaremos a mesma tabela usada na situação anterior para calcular o desvio médio, desvio padrão e variância, baseado.

Nela, basta selecionarmos a célula B4 e inserirmos o código:

$$= \text{DESV.MÉDIO}(B9 : B23)$$

Já para o cálculo desvio padrão, utilizamos um código parecido, inserindo ele na célula B5.

$$= \text{DESVPAD}(B9 : B23)$$

Figura 4.11: Cálculo do desvio padrão

	A	B
1	MÉDIA	76,4
2	MODA	80
3	MEDIANA	77
4	DESVIO MÉDIO	4,24
5	DESVIO PADRÃO	=DESVPAD(B9:B23)
6	VARIÂNCIA	31,83

Por fim, sabemos que a variância é o quadrado do desvio padrão, então para calcularmos, basta aplicarmos a potenciação no resultado anterior. Selecionando a B6, vamos inserir o código:

$$= \text{POTÊNCIA}(B5; 2)$$

A B5 contém o desvio padrão e o 2 representando que o número está elevado a 2 (ao quadrado), com isso obtemos a variância do conjunto, como expresso na Figura 4.12.

Figura 4.12: Cálculo da variância

	A	B
1	MÉDIA	76,4
2	MODA	80
3	MEDIANA	77
4	DESVIO MÉDIO	4,24
5	DESVIO PADRÃO	5,64
6	VARIÂNCIA	=POTÊNCIA(B5;2)

Com todas essas utilidades que o Microsoft Excel nos fornece, podemos enriquecer qualquer trabalho, com gráficos e tabelas, usando os cálculos para expressar os resultados das variáveis, que utilizamos para comparar, analisar e entender o comportamento da situação estudada.

Capítulo 5

Conclusão

O trabalho abordou a Estatística como ferramenta de comparação, servindo como guia para os professores e alunos do Ensino Secundário, tendo como referência teórica livros e trabalhos de autores da estatística. Exploramos o contexto futebolístico, ao usar dados sobre jogos, transformando-os em gráficos e tabelas, que foram apresentados e discutidos, por sua vez, como um resultado.

Contudo, o desenvolvimento do trabalho também sofreu influência da Pandemia enfrentada. Esse novo contexto trouxe como consequência vários desafios para os professores, como aulas online, aulas híbridas, gravações, planejamentos e estudos sobre novas didáticas, o que intensificou uma maior carga já existente, fazendo com que leituras e escritas sofressem atrasos. A partir disso, tivemos a principal dificuldade enfrentada na escrita do presente trabalho.

O desenvolvimento foi feito de modo detalhado, no qual buscamos dados e informações sobre os resultados dos clubes durante as 15 temporadas avaliadas. Essa coleta foi realizada no site da CBF e outros, com o objetivo de apresentar esses números não somente em uma tabela, como também em gráficos. Os cálculos de média foram resolvidos utilizando a ferramenta do Excel, como já mencionado, que, por sua vez, também buscamos expor sua utilização em outras ocasiões, tanto acadêmicas, profissionais ou pessoais.

Como já assinalado, desejávamos que o trabalho trouxesse conceitos estatísticos, a partir do esporte, mais especificamente o Futebol, como uma forma de apresentar a estatística através de suas ferramentas para exibir valores e resultados, por meio de tabelas, gráficos e medidas de posição e variação, a partir do Excel.

A apresentação foi feita com o objetivo de comparar o desempenho dos clubes do campeonato brasileiro , tanto campeões, como rebaixados, os G4, e outros grupos, por meio das pontuações, vitórias, empates, derrotas, gols marcados e sofridos, desempenho em campo de mando e em campo adversário, entre outras variáveis. Todo esse propósito foi atingido no desenvolver da dissertação, pois buscamos abordar com clareza todos os itens da estatística, ao apresentarmos todos os dados colhidos por meio de tabelas e gráficos. Além disso, comparando as informações ao buscarmos obter resultados padrões que serviram como meta para competições futuras.

Esperamos que o trabalho possa ter potencial utilização pelos professores do Ensino Secundário, ao aplicá-lo em sala de aula com seus alunos, adaptando-o às possíveis alterações desejadas pelo professor, tal como é usado em outro esporte. Desejamos também que o trabalho possa ser visualizado pelos alunos como um instrumento de direcionamento para outros que envolvam tanto a estatística, quanto a comparação e Excel, a fim de garantir uma aprendizagem mais significativa.

Portanto, esse trabalho poderá ser incrementado no futuro, com possibilidades de ser executado dentro de uma aplicação mais usual da estatística esportiva, como um trabalho voltado para avaliação de desempenho de jogadores (quando aplicado dentro de um esporte amador), com o objetivo de gerar uma ferramenta de apoio para o desenvolvimento de uma determinada equipe. Isso faz com que o trabalho possa ser utilizado tanto no âmbito esportivo, quanto no escolar.

Referências Bibliográficas

- [1] BERNARDO, Nívio. **Vamos estudar Física: introdução à Física**; Recife: Ed. do autor, 2004.
- [2] BUSSAB, Wilton de O.; MORETIIN, Pedro A. **Estatística Básica**. São Paulo: Saraiva, 2010.
- [3] CBF, Confederação Brasileira de Futebol. **Tabela do Campeonato Brasileiro de Futebol de 2012 à 2020 da Série A** . Disponível em: <https://www.cbf.com.br/futebol-brasileiro/competicoes/campeonato-brasileiro-serie-a/2020>; Acesso em 08/04/2021.
- [4] CORREA, Sonia Maria Barros Barbosa. **Probabilidade e Estatística** . Belo Horizonte: PUC Minas Virtual, 2003.
- [5] R. Bonita, R. Beaglehole, T. Kjellström. **Epidemiologia básica**. Santos, SP, 2010.
- [6] Época Negócios. **Como o Liverpool usou tecnologia e dados para vencer a Champions League**. Disponível em: <https://epocanegocios.globo.com/Empresa/noticia/2019/06/como-o-liverpool-usou-dados-e-tecnologia-para-chegar-final-da-champions-league.html>. Acesso em 30/05/2021.
- [7] ESPN **Resultados - Palmeiras** . Disponível em: <https://www.espn.com.br/futebol/time/resultados/id/2029/palmeiras>. Acesso em 10/04/2021.
- [8] FONSECA, Melissa L. **Excel básico**. Apostila Consultoria e Educação Corporativa. Curitiba, UFPR.

- [9] FERRARI, Fabrício. **Estatística Básica**. Porto Alegre: UFRGS, 2004. Classificação da série A. Disponível em: <https://www.flashscore.com.br/futebol/brasil/serie-a-2020/classificacao/>. Acesso em 21/04/2021.
- [10] FlashScore. **Classificação da série A 2020**. Disponível em: <https://www.flashscore.com.br/futebol/brasil/serie-a-2020/classificacao/>. Acesso em 03/04/2021.
- [11] Globo Esporte. **Balanco do Brasileirão 2018**. Disponível em: <https://interativos.globoesporte.globo.com/futebol/brasileirao-serie-a/especial/balanco-do-brasileirao-2018>.
- [12] GUEDES, Terezinha A.; MARTINS, Ana B. T.; ACORSI, Clédina R. L.; JANEIRO, Vanderly. **Aprender Fazendo Estatística**. São Paulo: Each USP, 2005.
- [13] HighBond. **Gráfico de área**. Disponível em: https://help.highbond.com/helpdocs/highbond/pt-br/Content/visualizations/interpretations/charts/stacked_area_chart.html. Acesso em 20 de junho de 2021.
- [14] LEVINE, David M.; BRENSON, Mark L.; Stephan, David. **Estatística: Teoria e aplicações usando o Microsoft Excel em Português**. LTC: City University of New York, 2017.
- [15] Liga Insights. **Os números que mudam o jogo: a influência dos dados no esporte**. Disponível em: <https://insights.liga.ventures/sports-techs/dados-no-esporte/>. Acesso em: 06/06/2021.
- [16] MARQUES, Mabel; GUIMARÃES, Gilda; GITIRANA, Verônica. **Compreensões de Alunos e Professores sobre Média Aritmética**. UFPE, 2011.
- [17] OPER. Como a estatística pode contribuir com o esporte?. Disponível em: <https://operdata.com.br/blog/estatistica-no-esporte/>. Acesso em: 31/05/2021.

- [18] Quero Apostar. **A importância da estatística para analisar o esporte americano**. Disponível em: <https://www.queroapostar.com/artigos/importancia-das-estatisticas-para-analisar-os-esportes-americanos/>. Acesso em: 31/05/2021.
- [19] RIBEIRO, Rafael do Nascimento Tomé. **Probabilidade e Estatística aplicada ao esporte**. Mossoró: UFERSA, 2016.
- [20] SALSA, Ivone da S.; MOREIRA, Janete A.; PEREIRA, Marcelo G. **Matemática e realidade: interdisciplinar**. Natal: EDUFRN, 2005.
- [21] SILVA, Michele dos S.; LIMA, Roberta da S. N.; VIEIRA, André R. L. **Excel como recurso didático nas aulas de estatística**. Senhor do Bonfim, UNEB.
- [22] SporTV. **Balanco do Brasileirão 2019**. Disponível em: <https://interativos.sportv.globo.com/site/especial/ep/balanco-do-brasileirao-2019>.
- [23] SR. GOOOL. Futebol, Rankings, Estatísticas e Simuladores. **Tabela do Campeonato Brasileiro de Futebol de 2019 da Série A**. Disponível em: <https://www.srgool.com.br/classificacao/Brasileirao/Serie-A/2019?a=todosclassificacao>. Acesso em 08/04/2021.
- [24] UOL. **Brasileirão 2006 à 2011: tabela**. Disponível em: <https://www.uol.com.br/esporte/futebol/campeonatos/brasileiro/2011/serie-a/tabela-de-jogos/tabela-de-jogos-1a-rodada.jhtm>. Acesso em 08/04/2021.
- [25] ZAT, Ancilla Dall'Onder. **Moda Estatística: Relações conceituais**. PUC-RS.
- [26] WHEELAN, C. **Estatística: o que é, para que serve, como funciona**. Rio de Janeiro: Zahar, 2016.

Apêndice

Figura 1: Tabela do campeonato brasileiro de 2006

#	Equipe	J	V	E	D	G	P
1.	 São Paulo	38	22	12	4	66:32	78
2.	 Internacional	38	20	9	9	52:36	69
3.	 Grêmio	38	20	7	11	64:45	67
4.	 Santos	38	18	10	10	58:36	64
5.	 Paraná	38	18	6	14	56:49	60
6.	 Vasco	38	15	14	9	57:50	59
7.	 Figueirense	38	15	12	11	52:44	57
8.	 Goiás	38	15	10	13	63:49	55
9.	 Cruzeiro	38	14	11	13	52:45	53
10.	 Corinthians	38	15	8	15	41:46	53
11.	 Flamengo	38	15	7	16	44:48	52
12.	 Botafogo	38	13	12	13	52:50	51
13.	 Athletico-PR	38	13	9	16	61:62	48
14.	 Juventude	38	13	8	17	44:54	47
15.	 Fluminense	38	11	12	15	48:58	45
16.	 Palmeiras	38	12	8	18	58:70	44
17.	 Ponte Preta	38	10	9	19	45:65	39
18.	 Fortaleza	38	8	14	16	39:62	38
19.	 São Caetano	38	9	9	20	37:53	36
20.	 Santa Cruz	38	7	7	24	41:76	28

Fonte: FlashScore

Figura 2: Tabela do campeonato brasileiro de 2007

#	Equipe	J	V	E	D	G	P
1.	 São Paulo	38	23	8	7	55:19	77
2.	 Santos	38	19	5	14	57:47	62
3.	 Flamengo	38	17	10	11	55:49	61
4.	 Fluminense	38	16	13	9	57:39	61
5.	 Cruzeiro	38	18	6	14	73:58	60
6.	 Palmeiras	38	16	10	12	48:47	58
7.	 Grêmio	38	17	7	14	44:43	58
8.	 Atlético-MG	38	15	10	13	63:51	55
9.	 Botafogo	38	14	13	11	62:58	55
10.	 Vasco	38	15	9	14	58:47	54
11.	 Internacional	38	15	9	14	49:44	54
12.	 Athletico-PR	38	14	12	12	51:50	54
13.	 Figueirense	38	14	11	13	57:56	53
14.	 Sport	38	14	9	15	54:55	51
15.	 Náutico	38	14	7	17	66:63	49
16.	 Goiás	38	13	6	19	49:62	45
17.	 Corinthians	38	10	14	14	40:50	44
18.	 Juventude	38	11	8	19	43:65	41
19.	 Paraná	38	11	8	19	42:64	41
20.	 América-RN	38	4	5	29	24:80	17

Fonte: FlashScore

Figura 3: Tabela do campeonato brasileiro de 2008

#	Equipe	J	V	E	D	G	P
1.	 São Paulo	38	21	12	5	66:36	75
2.	 Grêmio	38	21	9	8	59:35	72
3.	 Cruzeiro	38	21	4	13	59:44	67
4.	 Palmeiras	38	19	8	11	55:45	65
5.	 Flamengo	38	18	10	10	67:48	64
6.	 Internacional	38	15	9	14	48:47	54
7.	 Botafogo	38	15	8	15	51:44	53
8.	 Goiás	38	14	11	13	57:47	53
9.	 Coritiba	38	14	11	13	55:48	53
10.	 Vitória	38	15	7	16	48:44	52
11.	 Sport	38	14	10	14	48:45	52
12.	 Atlético-MG	38	12	12	14	50:61	48
13.	 Athletico-PR	38	12	9	17	45:54	45
14.	 Fluminense	38	11	12	15	49:48	45
15.	 Santos	38	11	12	15	44:53	45
16.	 Náutico	38	11	11	16	44:54	44
17.	 Figueirense	38	11	11	16	49:73	44
18.	 Vasco	38	11	7	20	56:72	40
19.	 Portuguesa	38	9	11	18	48:70	38
20.	 Ipatinga	38	9	8	21	37:67	35

Fonte: FlashScore

Figura 4: Tabela do campeonato brasileiro de 2009

#	Equipe	J	V	E	D	G	P
1.	 Flamengo	38	19	10	9	58:44	67
2.	 Internacional	38	19	8	11	65:44	65
3.	 São Paulo	38	18	11	9	57:42	65
4.	 Cruzeiro	38	18	8	12	58:53	62
5.	 Palmeiras	38	17	11	10	58:45	62
6.	 Avaí	38	15	12	11	61:52	57
7.	 Atlético-MG	38	16	8	14	55:56	56
8.	 Grêmio	38	15	10	13	67:46	55
9.	 Goiás	38	15	10	13	64:65	55
10.	 Corinthians	38	14	10	14	50:54	52
11.	 Barueri	38	12	13	13	59:52	49
12.	 Santos	38	12	13	13	58:58	49
13.	 Vitória	38	13	9	16	51:57	48
14.	 Athletico-PR	38	13	9	16	42:49	48
15.	 Botafogo	38	11	14	13	52:58	47
16.	 Fluminense	38	11	13	14	49:56	46
17.	 Coritiba	38	12	9	17	48:60	45
18.	 Santo André	38	11	8	19	46:61	41
19.	 Náutico	38	10	8	20	48:71	38
20.	 Sport	38	7	10	21	48:71	31

Fonte: FlashScore

Figura 5: Tabela do campeonato brasileiro de 2010

#	Equipe	J	V	E	D	G	P
1.	 Fluminense	38	20	11	7	62:36	71
2.	 Cruzeiro	38	20	9	9	53:38	69
3.	 Corinthians	38	19	11	8	65:41	68
4.	 Grêmio	38	17	12	9	68:43	63
5.	 Athletico-PR	38	17	9	12	43:45	60
6.	 Botafogo	38	14	17	7	54:42	59
7.	 Internacional	38	16	10	12	48:41	58
8.	 Santos	38	15	11	12	63:50	56
9.	 São Paulo	38	15	10	13	54:54	55
10.	 Palmeiras	38	12	14	12	42:43	50
11.	 Vasco	38	11	16	11	43:45	49
12.	 Ceará	38	10	17	11	35:44	47
13.	 Atlético-MG	38	13	6	19	52:64	45
14.	 Flamengo	38	9	17	12	41:44	44
15.	 Avaí	38	11	10	17	49:58	43
16.	 Atlético-GO	38	11	9	18	51:57	42
17.	 Vitória	38	9	15	14	42:48	42
18.	 Guarani	38	8	13	17	33:53	37
19.	 Goiás	38	8	9	21	41:68	33
20.	 Barueri	38	7	10	21	39:64	<u>28</u>

Fonte: FlashScore

Figura 6: Tabela do campeonato brasileiro de 2011

#	Equipe	J	V	E	D	G	P
1.	 Corinthians	38	21	8	9	53:36	71
2.	 Vasco	38	19	12	7	57:40	69
3.	 Fluminense	38	20	3	15	60:51	63
4.	 Flamengo	38	15	16	7	59:47	61
5.	 Internacional	38	16	12	10	57:43	60
6.	 São Paulo	38	16	11	11	57:46	59
7.	 Figueirense	38	15	13	10	46:45	58
8.	 Coritiba	38	16	9	13	57:41	57
9.	 Botafogo	38	16	8	14	52:49	56
10.	 Santos	38	15	8	15	55:55	53
11.	 Palmeiras	38	11	17	10	43:39	50
12.	 Grêmio	38	13	9	16	49:57	48
13.	 Atlético-GO	38	12	12	14	50:45	48
14.	 Bahia	38	11	13	14	43:49	46
15.	 Atlético-MG	38	13	6	19	50:60	45
16.	 Cruzeiro	38	11	10	17	48:51	43
17.	 Athletico-PR	38	10	11	17	38:55	41
18.	 Ceará	38	10	9	19	47:64	39
19.	 América-MG	38	8	13	17	51:69	37
20.	 Avai	38	7	10	21	45:75	31

Fonte: FlashScore

Figura 7: Tabela do campeonato brasileiro de 2012

#	Equipe	J	V	E	D	G	P
1.	 Fluminense	38	22	11	5	61:33	77
2.	 Atlético-MG	38	20	12	6	64:37	72
3.	 Grêmio	38	20	11	7	56:33	71
4.	 São Paulo	38	20	6	12	59:37	66
5.	 Vasco	38	16	10	12	45:44	58
6.	 Corinthians	38	15	12	11	51:39	57
7.	 Botafogo	38	15	10	13	60:50	55
8.	 Santos	38	13	14	11	50:44	53
9.	 Cruzeiro	38	15	7	16	47:51	52
10.	 Internacional	38	13	13	12	44:40	52
11.	 Flamengo	38	12	14	12	39:46	50
12.	 Náutico	38	14	7	17	44:51	49
13.	 Coritiba	38	14	6	18	53:60	48
14.	 Ponte Preta	38	12	12	14	37:44	48
15.	 Bahia	38	11	14	13	37:41	47
16.	 Portuguesa	38	10	15	13	39:41	45
17.	 Sport	38	10	11	17	39:56	41
18.	 Palmeiras	38	9	7	22	39:54	34
19.	 Atlético-GO	38	7	9	22	37:67	30
20.	 Figueirense	38	7	9	22	39:72	30

Fonte: FlashScore

Figura 8: Tabela do campeonato brasileiro de 2013

#	Equipe	J	V	E	D	G	P
1.	 Cruzeiro	38	23	7	8	77:37	76
2.	 Grêmio	38	18	11	9	42:35	65
3.	 Athletico-PR	38	18	10	10	65:49	64
4.	 Botafogo	38	17	10	11	55:41	61
5.	 Vitória	38	16	11	11	59:53	59
6.	 Goiás	38	16	11	11	48:44	59
7.	 Santos	38	15	12	11	51:38	57
8.	 Atlético-MG	38	15	12	11	49:38	57
9.	 São Paulo	38	14	8	16	39:40	50
10.	 Corinthians	38	11	17	10	27:22	50
11.	 Coritiba	38	12	12	14	42:45	48
12.	 Bahia	38	12	12	14	37:45	48
13.	 Internacional	38	11	15	12	51:52	48
14.	 Criciúma	38	13	7	18	49:63	46
15.	 Fluminense	38	12	10	16	43:47	46
16.	 Flamengo	38	12	13	13	43:46	45
17.	 Portuguesa	38	12	12	14	50:46	44
18.	 Vasco	38	11	11	16	50:61	44
19.	 Ponte Preta	38	9	10	19	37:55	37
20.	 Náutico	38	5	5	28	22:79	20

Fonte: FlashScore

Figura 9: Tabela do campeonato brasileiro de 2014

#	Equipe	J	V	E	D	G	P
1.	 Cruzeiro	38	24	8	6	67:38	80
2.	 São Paulo	38	20	10	8	59:40	70
3.	 Internacional	38	21	6	11	53:41	69
4.	 Corinthians	38	19	12	7	49:31	69
5.	 Atlético-MG	38	17	11	10	51:38	62
6.	 Fluminense	38	17	10	11	61:42	61
7.	 Grêmio	38	17	10	11	36:24	61
8.	 Athletico-PR	38	15	9	14	43:42	54
9.	 Santos	38	15	8	15	42:35	53
10.	 Flamengo	38	14	10	14	46:47	52
11.	 Sport	38	14	10	14	36:46	52
12.	 Goiás	38	13	8	17	38:40	47
13.	 Figueirense	38	13	8	17	37:47	47
14.	 Coritiba	38	12	11	15	42:45	47
15.	 Chapecoense	38	11	10	17	39:44	43
16.	 Palmeiras	38	11	7	20	34:59	40
17.	 Vitória	38	10	8	20	37:54	38
18.	 Bahia	38	9	10	19	31:43	37
19.	 Botafogo	38	9	7	22	31:48	34
20.	 Criciúma	38	7	11	20	28:56	32

Fonte: FlashScore

Figura 10: Tabela do campeonato brasileiro de 2015

#	Equipe	J	V	E	D	G	P
1.	 Corinthians 	38	24	9	5	71:31	81
2.	 Atlético-MG 	38	21	6	11	65:47	69
3.	 Grêmio 	38	20	8	10	52:32	68
4.	 São Paulo	38	18	8	12	53:47	62
5.	 Internacional	38	17	9	12	39:38	60
6.	 Sport	38	15	14	9	53:38	59
7.	 Santos	38	16	10	12	59:41	58
8.	 Cruzeiro	38	15	10	13	44:35	55
9.	 Palmeiras	38	15	8	15	60:51	53
10.	 Athletico-PR	38	14	9	15	43:48	51
11.	 Ponte Preta	38	13	12	13	41:40	51
12.	 Flamengo	38	15	4	19	45:53	49
13.	 Fluminense	38	14	5	19	40:49	47
14.	 Chapecoense	38	12	11	15	34:44	47
15.	 Coritiba	38	11	11	16	31:42	44
16.	 Figueirense	38	11	10	17	36:50	43
17.	 Avaí	38	11	9	18	38:60	42
18.	 Vasco	38	10	11	17	28:54	41
19.	 Goiás 	38	10	8	20	39:49	38
20.	 Joinville 	38	7	10	21	26:48	31

Fonte: FlashScore

Figura 11: Tabela do campeonato brasileiro de 2016

#	Equipe	J	V	E	D	G	P
1.	 Palmeiras	38	24	8	6	62:32	80
2.	 Santos	38	22	5	11	59:35	71
3.	 Flamengo	38	20	11	7	52:35	71
4.	 Atlético-MG	38	17	11	10	61:53	62
5.	 Botafogo	38	17	8	13	43:39	59
6.	 Athletico-PR	38	17	6	15	38:32	57
7.	 Corinthians	38	15	10	13	48:42	55
8.	 Ponte Preta	38	15	8	15	48:52	53
9.	 Grêmio	38	14	11	13	41:44	53
10.	 São Paulo	38	14	10	14	44:36	52
11.	 Chapecoense	38	13	13	12	49:56	52
12.	 Cruzeiro	38	14	9	15	48:49	51
13.	 Fluminense	38	13	11	14	45:45	50
14.	 Sport	38	13	8	17	49:55	47
15.	 Coritiba	38	11	13	14	41:42	46
16.	 Vitória	38	12	9	17	51:53	45
17.	 Internacional	38	11	10	17	35:41	43
18.	 Figueirense	38	8	13	17	30:50	37
19.	 Santa Cruz	38	8	7	23	45:69	31
20.	 América-MG	38	7	7	24	23:58	28

Fonte: FlashScore

Figura 12: Tabela do campeonato brasileiro de 2017

#	Equipe	J	V	E	D	G	P
1.	 Corinthians	38	21	9	8	50:30	72
2.	 Palmeiras	38	19	6	13	61:45	63
3.	 Santos	38	17	12	9	42:32	63
4.	 Grêmio	38	18	8	12	55:36	62
5.	 Cruzeiro	38	15	12	11	47:39	57
6.	 Flamengo	38	15	11	12	49:38	56
7.	 Vasco	38	15	11	12	40:47	56
8.	 Chapecoense	38	15	9	14	47:49	54
9.	 Atlético-MG	38	14	12	12	52:49	54
10.	 Botafogo	38	14	11	13	45:42	53
11.	 Athletico-PR	38	14	9	15	45:43	51
12.	 Bahia	38	13	11	14	50:48	50
13.	 São Paulo	38	13	11	14	48:49	50
14.	 Fluminense	38	11	14	13	50:53	47
15.	 Sport	38	12	9	17	46:58	45
16.	 Vitória	38	11	10	17	50:58	43
17.	 Coritiba	38	11	10	17	42:51	43
18.	 Avaí	38	10	13	15	29:48	43
19.	 Ponte Preta	38	10	9	19	37:52	39
20.	 Atlético-GO	38	9	9	20	38:56	36

Fonte: FlashScore

Figura 13: Tabela do campeonato brasileiro de 2018

#	Equipe	J	V	E	D	G	P
1.	 Palmeiras	38	23	11	4	64:26	80
2.	 Flamengo	38	21	9	8	59:29	72
3.	 Internacional	38	19	12	7	51:29	69
4.	 Grêmio	38	18	12	8	48:27	66
5.	 São Paulo	38	16	15	7	46:34	63
6.	 Atlético-MG	38	17	8	13	56:43	59
7.	 Athletico-PR	38	16	9	13	54:37	57
8.	 Cruzeiro	38	14	11	13	34:34	53
9.	 Botafogo	38	13	12	13	38:46	51
10.	 Santos	38	13	11	14	46:40	50
11.	 Bahia	38	12	12	14	39:41	48
12.	 Fluminense	38	12	9	17	32:46	45
13.	 Corinthians	38	11	11	16	34:35	44
14.	 Chapecoense	38	11	11	16	34:50	44
15.	 Ceará	38	10	14	14	32:38	44
16.	 Vasco	38	10	13	15	41:48	43
17.	 Sport	38	11	9	18	35:57	42
18.	 América-MG	38	10	10	18	30:47	40
19.	 Vitória	38	9	10	19	36:63	37
20.	 Paraná	38	4	11	23	18:57	23

Fonte: FlashScore

Figura 14: Tabela do campeonato brasileiro de 2019

#	Equipe	J	V	E	D	G	P
1.	 Flamengo	38	28	6	4	86:37	90
2.	 Santos	38	22	8	8	60:33	74
3.	 Palmeiras	38	21	11	6	61:32	74
4.	 Grêmio	38	19	8	11	64:39	65
5.	 Athletico-PR	38	18	10	10	51:32	64
6.	 São Paulo	38	17	12	9	39:30	63
7.	 Internacional	38	16	9	13	44:39	57
8.	 Corinthians	38	14	14	10	42:34	56
9.	 Fortaleza	38	15	8	15	50:49	53
10.	 Goiás	38	15	7	16	46:64	52
11.	 Bahia	38	12	13	13	44:43	49
12.	 Vasco	38	12	13	13	39:45	49
13.	 Atlético-MG	38	13	9	16	45:49	48
14.	 Fluminense	38	12	10	16	38:46	46
15.	 Botafogo	38	13	4	21	31:45	43
16.	 Ceará	38	10	9	19	36:41	39
17.	 Cruzeiro	38	7	15	16	27:46	36
18.	 CSA	38	8	8	22	24:58	32
19.	 Chapecoense	38	7	11	20	31:52	32
20.	 Avai	38	3	11	24	18:62	20

Fonte: FlashScore

Figura 15: Tabela do campeonato brasileiro de 2020

#	Equipe	J	V	E	D	G	P
1.	 Flamengo	38	21	8	9	68:48	71
2.	 Internacional	38	20	10	8	61:35	70
3.	 Atlético-MG	38	20	8	10	64:45	68
4.	 São Paulo	38	18	12	8	59:41	66
5.	 Fluminense	38	18	10	10	55:42	64
6.	 Grêmio	38	14	17	7	53:40	59
7.	 Palmeiras	38	15	13	10	51:37	58
8.	 Santos	38	14	12	12	52:51	54
9.	 Athletico-PR	38	15	8	15	38:36	53
10.	 Red Bull Bragantino	38	13	14	11	50:40	53
11.	 Ceará	38	14	10	14	54:51	52
12.	 Corinthians	38	13	12	13	45:45	51
13.	 Atlético-GO	38	12	14	12	40:45	50
14.	 Bahia	38	12	8	18	48:59	44
15.	 Sport	38	12	6	20	31:50	42
16.	 Fortaleza	38	10	11	17	34:44	41
17.	 Vasco	38	10	11	17	37:56	41
18.	 Goiás	38	9	10	19	41:63	37
19.	 Coritiba	38	7	10	21	31:54	31
20.	 Botafogo	38	5	12	21	32:62	27

Fonte: FlashScore