

Análise de dados esportivos

Premier League

João Fagundes

joao.230702@alunos.utfpr.edu.br

Universidade Tecnológica Federal do Paraná

Ponta Grossa, Paraná

1 Introdução

Este estudo foca no desempenho das equipes ao longo das temporadas da Premier League, utilizando análise de dados e programação em Python. O objetivo é identificar padrões de desempenho distintos entre as equipes, oferecendo insights para gestores esportivos e analistas. A análise se baseia em métricas como Expected Goals (xG), Expected Goals Against (xGA) e pontos acumulados, utilizando técnicas avançadas como clustering com o algoritmo KMeans.

2 Extração de dados

Para começar, o processo de extração de dados foi realizado usando Python e várias bibliotecas úteis para web scraping e manipulação de dados, como requests, BeautifulSoup e pandas. O objetivo era obter informações detalhadas sobre jogos e estatísticas da Premier League do site Understat.

2.1 Coleta do IDs dos jogos

Os IDs dos jogos da temporada 2024 da Premier League foram coletados do site Understat. Os dados foram extraídos de um script JSON na página, decodificados e convertidos em um DataFrame do Pandas para análises futuras.

2.2 Coleta de estatísticas dos jogos

As estatísticas detalhadas de cada jogo da temporada 2024 da Premier League foram extraídas do site Understat. Iterando sobre a lista de IDs de jogos previamente obtida, uma solicitação foi feita para cada URL específica de jogo (https://understat.com/match/{id_do_jogo}). Utilizando expressões regulares, os dados JSON relevantes foram extraídos diretamente das páginas HTML de cada jogo. Após decodificação e processamento, esses dados foram transformados em DataFrames do Pandas e concatenados para formar um conjunto abrangente de estatísticas de todos os jogos analisados.

2.3 Coleta dos tweets

Para coletar os tweets usei um dataset no formato csv com todas as informações necessárias para realizar a análise.

2.4 Extração das tabelas de temporadas anteriores da Premier League

Para analisar o desempenho histórico das equipes na Premier League, foram extraídas as tabelas de várias temporadas. O processo envolveu acessar páginas específicas do site Understat para cada temporada (2024, 2022, 2020, 2018, 2016, 2014, 2012). Utilizando Python, os dados foram coletados, processados e transformados em estruturas de dados tabulares usando a biblioteca Pandas. As estatísticas agregadas, como vitórias, empates, derrotas, gols marcados e sofridos, foram calculadas para cada temporada. Isso proporcionou uma visão comparativa do desempenho das equipes ao longo do tempo, destacando padrões e tendências na Premier League.

2.5 Coleta de estatísticas dos jogadores

Para extrair os dados dos jogadores da Premier League, foi realizada uma requisição HTTP para acessar a página principal da liga no site Understat. Ao analisar o conteúdo da página com BeautifulSoup, identificou-se o script que continha as informações dos jogadores, verificando a presença da substring "playersData".

Uma vez localizado o script relevante, os dados JSON foram extraídos e decodificados para garantir a correta interpretação de caracteres especiais. Em seguida, esses dados foram estruturados em um DataFrame do Pandas, proporcionando uma organização tabular para facilitar a manipulação e análise subsequentes.

2.6 Construção da tabela de classificação da Liga

Um processo automatizado foi implementado para manter a tabela da Premier League atualizada com base nos resultados dos jogos. Inicialmente, a estrutura da tabela é definida, incluindo campos como equipe, partidas jogadas, vitórias, empates, derrotas, gols marcados, gols sofridos, diferença de gols e pontos acumulados. Após cada jogo, as estatísticas de cada equipe são ajustadas conforme necessário, registrando partidas jogadas, gols marcados e sofridos, calculando a diferença de gols e atribuindo pontos com base no resultado (vitória, empate ou derrota). Equipes são adicionadas automaticamente se não estiverem listadas, começando com estatísticas zeradas. Após a atualização dos resultados, a tabela é ordenada por pontos acumulados e, em seguida, por diferença de gols, garantindo uma classificação precisa das equipes de acordo com seu desempenho na liga.

	team	played	won	drawn	lost	goals_for	goals_against	goal_difference	points
0	Manchester City	38	28	7	3	96	34	62	91
1	Arsenal	38	28	5	5	91	29	62	89
2	Liverpool	38	24	10	4	86	41	45	82
3	Aston Villa	38	20	8	10	76	61	15	68
4	Tottenham	38	20	6	12	74	61	13	66
5	Chelsea	38	18	9	11	77	63	14	63
6	Newcastle United	38	18	6	14	85	62	23	60
7	Manchester United	38	18	6	14	57	58	-1	60
8	West Ham	38	14	10	14	60	74	-14	52
9	Crystal Palace	38	13	10	15	57	58	-1	49
10	Brighton	38	12	12	14	55	62	-7	48
11	Everton	38	13	9	16	40	51	-11	48
12	Bournemouth	38	13	9	16	54	67	-13	48
13	Fulham	38	13	8	17	55	61	-6	47
14	Wolverhampton Wanderers	38	13	7	18	50	65	-15	46
15	Brentford	38	10	9	19	56	65	-9	39
16	Nottingham Forest	38	9	9	20	49	67	-18	36
17	Luton	38	6	8	24	52	85	-33	26
18	Burnley	38	5	9	24	41	78	-37	24
19	Sheffield United	38	3	7	28	35	104	-69	16

Figure 1. Tabela de Classificação da Premier League.

3 Análise estatística das equipes

Nesta seção, foi realizada uma análise detalhada das estatísticas das equipes da Premier League, utilizando dados da temporada. O objetivo é fornecer uma compreensão abrangente do desempenho de cada equipe em vários aspectos chave.

Para cada equipe, foi calculado e examinado métricas essenciais como gols marcados, gols sofridos, expected goals (xG), taxa de conversão de gols, precisão dos chutes no alvo e PPDA médio (Pressing Per Defensive Action). Essas métricas oferecem insights valiosos sobre a eficiência ofensiva, defensiva e tática de cada time ao longo dos jogos.

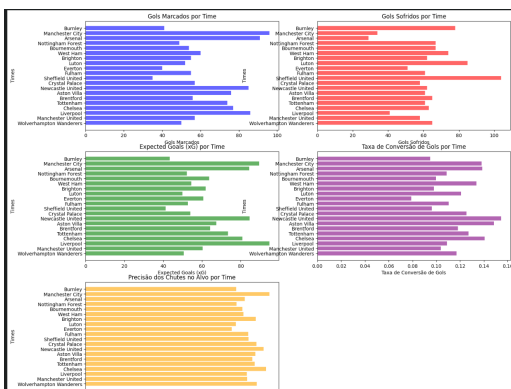


Figure 2. Análise comparativa entre os times

Manchester City lidera a Premier League com uma ofensiva altamente eficaz, marcando a maior taxa de gols por jogo e exibindo a melhor precisão de chutes no gol, além de um alto valor de Expected Goals (xG), refletindo sua capacidade de criar e capitalizar oportunidades de gol. Em contraste, o Sheffield United, último colocado, enfrenta desafios defensivos significativos, sofrendo a maior taxa de gols por jogo na liga e apresentando uma baixa taxa de Expected Goals (xG), indicando problemas na criação de chances de gol de qualidade. A baixa taxa de conversão de gols do Sheffield

United sugere dificuldades na finalização, destacando suas limitações frente ao ataque eficiente do Manchester City.

3.1 Informações adicionais sobre estatística

A tabela a seguir apresenta uma análise detalhada do desempenho das equipes da Premier League em termos de gols marcados, sofridos e taxas de vitória, empate e derrota em jogos realizados em casa e fora. Os dados foram processados utilizando um código especializado para calcular e comparar métricas fundamentais de cada equipe ao longo da temporada.

	Team	Home Goals	Home Wins	Home Draws	Home Losses	Home Goals Against	Home Points	Away Goals	Away Wins	Away Draws	Away Losses	Away Goals Against	Away Points	Total Goals	Total Wins	Total Draws	Total Losses	Total Points
0	Manchester City	52	10	5	3	19	30	26	10	5	13	21	15	78	20	10	13	50
1	Arsenal	49	10	5	3	16	28	23	9	4	11	18	13	72	14	9	14	47
2	Liverpool	46	10	6	2	25	26	20	8	4	12	27	16	66	14	10	14	48
3	Aston Villa	40	8	7	3	28	23	16	6	5	7	25	11	56	14	12	7	41
4	Tottenham	39	7	6	5	25	22	17	6	4	8	24	14	56	11	10	17	35
5	Chelsea	38	6	7	5	26	21	19	5	4	9	24	13	57	10	11	17	31
6	Newcastle United	45	8	6	4	20	26	18	5	3	10	25	13	63	11	9	14	33
7	Manchester United	31	6	6	6	27	24	26	4	3	11	31	11	57	9	9	16	28
8	West Ham	33	5	7	6	28	25	27	3	2	13	30	8	60	7	9	22	26
9	Crystal Palace	28	4	6	8	29	22	26	2	1	15	29	7	54	5	7	26	17
10	Brighton	30	4	6	8	26	20	24	2	1	15	26	7	54	5	7	26	17
11	Everton	22	3	5	10	28	15	18	1	1	16	24	4	40	4	6	20	10
12	Bournemouth	27	4	5	9	26	16	27	2	1	15	28	7	54	5	6	21	13
13	Fulham	28	4	6	8	27	16	29	2	1	15	28	7	57	5	7	22	12
14	Wolverhampton Wanderers	25	3	5	10	28	12	26	1	1	16	27	4	51	4	6	22	10
15	Brentford	30	3	6	9	25	14	28	1	1	16	27	4	58	4	7	23	11
16	Nottingham Forest	26	3	5	10	27	13	25	1	1	16	26	4	51	4	6	22	10
17	Luton	28	2	4	12	30	8	31	0	0	18	31	0	59	2	4	22	6
18	Burnley	21	2	4	12	29	10	28	0	0	18	29	0	49	2	4	22	6
19	Sheffield United	10	1	3	14	35	3	44	0	0	28	41	0	54	1	3	28	3

Figure 3. Análise comparativa entre os times

O Manchester City lidera a Premier League com excelentes médias de gols marcados tanto em casa (2.68) quanto fora (2.37), acompanhadas por altas taxas de vitória em ambas as condições (73.68%). Isso destaca sua eficiência tanto no ataque quanto na defesa, consolidando sua posição no topo da tabela. Em contraste, o Sheffield United enfrenta sérias dificuldades defensivas, sofrendo em média 2.47 gols por jogo fora de casa e tendo uma alta taxa de derrota de 78.95%, contribuindo para sua posição na parte inferior da classificação. O Arsenal exibe um desempenho forte em casa com uma média de 2.53 gols marcados e uma alta taxa de vitória de 78.95%. Liverpool e Tottenham mostram consistência em marcar gols em casa, enquanto Chelsea e Manchester United demonstram desempenho equilibrado em ambos os cenários de jogo. Por outro lado, o Luton enfrenta dificuldades significativas com baixas médias de gols marcados, altas médias de gols sofridos e altas taxas de derrota, refletindo seus desafios competitivos na liga.

3.2 Relação entre Baixa Sofrimento de Gols e Desempenho

Observando os dados até agora, podemos inferir que há uma relação clara entre a capacidade de um time de evitar gols e seu sucesso geral em partidas. Equipes como Arsenal e Manchester City, que sofrem poucos gols tanto em casa quanto fora, tendem a ter taxas mais altas de vitórias e menos derrotas. Isso sugere que uma defesa robusta não apenas reduz os riscos de derrota, mas também fornece uma base sólida para a conquista de pontos cruciais ao longo da temporada.

3.3 Eficiência em jogos fora de casa

Algumas equipes mostram uma notável capacidade de adaptar sua estratégia quando jogam fora de casa. Manchester

City, por exemplo, mantém uma taxa de vitória impressionante de 73.68% em jogos fora, indicando que sua abordagem é eficaz em diferentes ambientes e condições.

Outras equipes, como Tottenham e Liverpool, também demonstram um bom equilíbrio entre defesa e ataque, contribuindo para taxas de vitória consistentes tanto em casa quanto fora.

4 Análise do comportamento dos times

Também foi desenvolvido um código para calcular as médias de PPDA (Passes Allowed Per Defensive Action) para cada time, tanto em jogos realizados em casa quanto fora. Esse código utiliza operações simples de manipulação de dados para calcular e apresentar as médias de PPDA de cada equipe, proporcionando uma visão detalhada da eficiência defensiva de cada time em diferentes contextos de jogo.

Tabela de Médias de PPDA por Time:

	Team	Average Home PPDA	Average Away PPDA
0	Arsenal	9.953211	9.561068
1	Aston Villa	13.420884	18.010095
2	Bournemouth	9.940163	10.790716
3	Brentford	10.850989	11.135847
4	Brighton	10.680084	10.631800
5	Burnley	13.634663	14.934400
6	Chelsea	11.340695	9.730074
7	Crystal Palace	13.677068	16.429411
8	Everton	14.087774	16.015316
9	Fulham	15.853653	15.430005
10	Liverpool	6.806705	9.256916
11	Luton	11.775237	17.592095
12	Manchester City	11.321532	12.481942
13	Manchester United	11.883021	13.899889
14	Newcastle United	9.834842	12.148679
15	Nottingham Forest	20.190511	22.319358
16	Sheffield United	27.304147	22.593432
17	Tottenham	7.952995	7.709858
18	West Ham	17.741821	19.549353
19	Wolverhampton Wanderers	12.124584	14.972616

Figure 4. Análise de PPDA

4.1 Eficiência defensiva e PPDA

O Arsenal possui um Average Home PPDA de 9.95 e Average Away PPDA de 9.56. Isso indica que o Arsenal pressiona menos quando joga fora, o que pode estar correlacionado com seu estilo de jogo mais cauteloso fora de casa. Apesar disso, mantém uma das defesas mais sólidas, sofrendo apenas 0.84 gols em média em casa e 0.95 fora. Isso sugere que o Arsenal consegue manter uma boa organização defensiva, mesmo sem pressionar tanto o adversário fora de casa.

Já o Manchester City, possui um Average Home PPDA de 11.34 e Average Away PPDA de 12.48. Embora tenha uma pressão defensiva um pouco mais alta do que o Arsenal, ainda mantém uma defesa eficiente, sofrendo 0.84 gols em casa e 0.95 gols fora. Isso pode indicar uma estratégia de jogo que se adapta bem tanto em casa quanto fora, embora sua pressão defensiva possa variar ligeiramente dependendo do local da partida.

4.2 Eficiência defensiva e taxas de gols sofridos

O Liverpool apesar de ter um PPDA muito baixo em casa (6.81) e moderado fora (9.26), o Liverpool sofreu apenas 0.89 gols em média em casa, mas 1.26 gols fora. Isso sugere que, embora o Liverpool pressione mais em casa, sua eficácia defensiva pode ser comprometida em partidas fora de Anfield. Isso pode estar relacionado a uma estratégia mais agressiva em casa, contrastando com um possível desafio em manter a mesma solidez defensiva fora de casa.

4.3 Equipes com PPDA alto e defesas vulneráveis

Sheffield United possui o maior Average Home PPDA (27.30) e Average Away PPDA (22.59), o que indica uma defesa menos organizada e mais permissiva em ambas as situações. Isso se correlaciona com as altas taxas de gols sofridos em casa (3.00) e fora (2.47), refletindo uma vulnerabilidade defensiva significativa que contribui para suas altas taxas de derrota.

Já o Nottingham Forest possui um Average Home PPDA de 20.19 e Average Away PPDA de 22.32. Assim como o Sheffield United, apresenta uma defesa vulnerável, refletida nas altas médias de gols sofridos tanto em casa (1.58) quanto fora (1.95). Esses números indicam dificuldades consistentes na organização defensiva, independentemente do local da partida.

5 Informações sobre jogadores

Para compreender melhor o desempenho dos jogadores ao longo da temporada, foi desenvolvido um código para identificar os principais artilheiros com base no número de gols marcados. A seguir, apresento os resultados obtidos

id	player_name	goals	assists	minutes	goals_per_minute	assists_per_minute	goals_per_90	assists_per_90	goals_per_90_minus_assists_per_90
1	Erling Haaland	11	2	1120	0.0098	0.0018	1.08	0.18	0.90
2	Cole Palmer	11	1	1120	0.0098	0.0009	1.08	0.09	0.99
3	Alexander Isak	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
4	Phil Jones	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
5	Anthony Gordon	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
6	Chris Wood	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
7	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
8	Michael Okeke	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
9	Chris Wood	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
10	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
11	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
12	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
13	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
14	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
15	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
16	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
17	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
18	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
19	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90
20	Sam Kerr	10	1	1120	0.0089	0.0009	0.99	0.09	0.90

Figure 5. Jogadores com mais gols

A análise dos artilheiros da temporada, juntamente com os dados de PPDA (Passes Permitidos por Ação Defensiva) e a posição na tabela da Premier League, revela conexões significativas. Jogadores como Erling Haaland (Manchester City), Cole Palmer (Chelsea e Manchester City) e Alexander Isak (Newcastle United) destacam-se não apenas pelos gols marcados, mas também pela capacidade de criar jogadas e finalizar eficientemente. Eles contribuem não só com gols, mas também com assistências e passes-chave. O PPDA mais baixo de certas equipes indica uma pressão defensiva intensa, ajudando a recuperar a posse de bola em áreas avançadas do campo e a criar mais oportunidades de gol. Equipes bem posicionadas na tabela combinam um forte desempenho ofensivo com uma defesa organizada, refletindo consistência ao longo da temporada. A presença de artilheiros entre os líderes de

gols está diretamente relacionada à posição elevada na classificação, destacando o impacto decisivo do desempenho individual dos jogadores no sucesso coletivo da equipe.

5.1 K-Means: Clusterização dos jogadores

Além da análise dos principais artilheiros e sua influência nos resultados, utilizei o algoritmo K-means para agrupar os jogadores em clusters com base em suas estatísticas de desempenho. As features consideradas foram gols marcados, expected goals (xG), assistências, expected assists (xA), número de chutes a gol e passes-chave.

Estatísticas por Cluster:									
cluster	goals				assists				
	mean	std	min	max	mean	std	min	max	
0	4.727273	2.582262	0	10	6.818182	2.223267	2	10	
1	0.455959	0.775855	0	5	0.409326	0.744375	0	3	
2	14.583333	4.959985	7	27	6.541667	2.918091	2	13	
3	3.982759	2.895222	0	14	2.086207	1.374102	0	6	

Figure 6. Clusters

Os dados revelam quatro clusters distintos de jogadores na análise. No Cluster 0, os jogadores apresentam uma média de 4.73 gols e 6.82 assistências, com um desvio padrão de 2.58, indicando consistência razoável tanto na capacidade de marcar quanto na criação de jogadas. No Cluster 1, os números são significativamente mais baixos, com uma média de apenas 0.46 gols e 0.41 assistências, sugerindo um papel limitado tanto na finalização quanto na criação de oportunidades. O Cluster 2 destaca-se com uma média elevada de 14.58 gols e 6.54 assistências, mas com uma grande variação (desvio padrão de 4.96), indicando jogadores excepcionalmente prolíficos tanto na pontuação quanto na criação de jogadas. Por fim, o Cluster 3 mostra uma média de 3.98 gols e 2.08 assistências, refletindo contribuições mais modestas tanto em gols quanto na criação de oportunidades. Esses clusters oferecem insights sobre como diferentes clubes distribuem seus jogadores entre diferentes perfis de desempenho, o que influencia diretamente a estratégia e o estilo de jogo de suas equipes.

Times que mais aparecem em cada cluster:									
Cluster 0:									
['Liverpool', 'Aston Villa', 'Manchester City', 'West Ham', 'Arsenal']									
Cluster 1:									
['Sheffield United', 'Nottingham Forest', 'Aston Villa', 'Chelsea', 'Brighton']									
Cluster 2:									
['Liverpool', 'Manchester City', 'Arsenal', 'Crystal Palace', 'Manchester United']									
Cluster 3:									
['Burnley', 'Bournemouth', 'Everton', 'Fulham', 'Brighton']									

Figure 7. Clusters por time

Os dados revelam quatro clusters distintos de times na análise. No Cluster 0, incluindo Liverpool, Manchester City e Arsenal, os times destacam-se por jogadores que marcam gols consistentemente e contribuem na criação de jogadas para os colegas, adotando um estilo de jogo ofensivo. No Cluster 1, clubes como Sheffield United e Nottingham Forest

dependem mais de um coletivo sólido e enfrentam desafios na criação de oportunidades. O Cluster 2 é liderado por Liverpool e Manchester City, com ênfase em finalização eficaz e conversão de oportunidades em gols, enquanto no Cluster 3, clubes como Burnley e Bournemouth mostram uma distribuição equilibrada de responsabilidades ofensivas e um jogo mais cauteloso. Esses clusters revelam como diferentes estratégias táticas e estilos de jogo influenciam a composição e o desempenho dos times na Premier League.

6 Desempenho ao longo dos anos

Foi elaborado um gráfico de linha para visualizar de a evolução dos pontos dos times ao longo dos anos selecionados. Utilizamos os dados disponíveis, plotando a variação dos pontos acumulados em cada temporada ao longo do tempo. O gráfico permite uma análise direta da performance desses times, destacando tendências e variações significativas ao longo dos anos estudados. A figura mostra um exemplo que selecionando alguns dos times disponíveis.

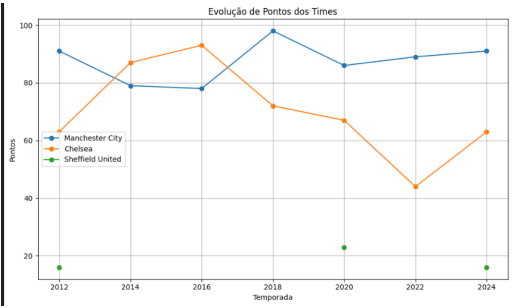


Figure 8. Desempenho ao longo das temporadas

O Sheffield United aparece nos dados em temporadas específicas devido ao seu rebaixamento para divisões inferiores em anos intercalados, resultando na ausência da tabela principal da liga em certos períodos. O Chelsea mostra uma variação significativa na pontuação ao longo das temporadas devido a diversos fatores internos e externos, incluindo mudanças na direção técnica e gestão. Em contraste, o Manchester City demonstra maior consistência nos resultados ao longo das temporadas, especialmente em métricas como PPDA, sugerindo uma defesa organizada e eficaz.

6.1 Entendendo o desempenho do Sheffield United

O Sheffield United frequentemente ocupou posições inferiores na tabela classificativa nas temporadas em que esteve presente devido ao seu histórico de luta contra o rebaixamento. O clube não se destacou por alta taxa de gols marcados ou por uma defesa excepcional, resultando em saldo de gols desfavorável e menor acumulação de pontos. Análises de métricas como PPDA sugerem que o estilo de jogo do clube não favorece consistentemente a obtenção de pontos

ao longo das temporadas. Fatores como eficiência na criação de chances de gol e conversão dessas chances também influenciam diretamente o desempenho do clube.

7 K-Means: Aplicando Clusters Nos Times

Utilizando o algoritmo de aprendizado não supervisionado K-means, normalizei métricas como Expected Goals (xG), Expected Goals Against (xGA) e pontos (pts) usando StandardScaler para garantir consistência na escala. Apliquei o K-means para agrupar os times da Premier League em 4 clusters distintos com base nessas características normalizadas. Fiz isso para analisar a dinâmica dos times ao longo do tempo, especialmente focando em Manchester United, Manchester City e Chelsea. Plotei gráficos que ilustram como esses times variaram entre os clusters ao longo das temporadas, proporcionando insights sobre a evolução do desempenho e possíveis mudanças na performance ao longo do tempo. Essa abordagem permite visualizar claramente como os times foram agrupados e como esses agrupamentos mudaram, revelando padrões de consistência ou variação no desempenho ao longo das observações.

```
Informações dos Clusters:
Cluster 0:
xG      63.614260
xGA     56.039251
pts     59.906977
Name: 0, dtype: float64

Cluster 1:
xG      44.485915
xGA     55.144058
pts     43.163265
Name: 1, dtype: float64

Cluster 2:
xG      43.320590
xGA     73.561708
pts     32.080000
Name: 2, dtype: float64

Cluster 3:
xG      77.049244
xGA     35.633956
pts     82.826087
Name: 3, dtype: float64
```

Figure 9. Clusters dos times

Utilizei o algoritmo K-means para agrupar times da Premier League com base em métricas como Expected Goals

(xG), Expected Goals Against (xGA) e pontos (pts), normalizadas com StandardScaler para garantir consistência na escala. Focando em Manchester United, Manchester City e Chelsea, analisei como esses times variaram entre os 4 clusters ao longo das temporadas, plotando gráficos para visualizar mudanças na performance e identificar padrões de consistência ou variação ao longo do tempo

7.1 Gráficos Dos Clusters

No contexto da análise de clusters realizada, o Manchester City demonstrou uma notável consistência ao longo dos anos citados, mantendo-se predominantemente no Cluster 3. Este cluster é caracterizado por um alto valor médio de Expected Goals (xG), uma defesa robusta com baixo Expected Goals Against (xGA) e uma acumulação elevada de pontos. Esses resultados indicam que o Manchester City tem sido consistentemente eficaz na criação de chances de gol, sólida defesa e excelente desempenho na acumulação de pontos ao longo das temporadas analisadas. Essa estabilidade no Cluster 3 sugere uma performance geral muito positiva e uma capacidade de manter um padrão elevado de desempenho ao longo do tempo.

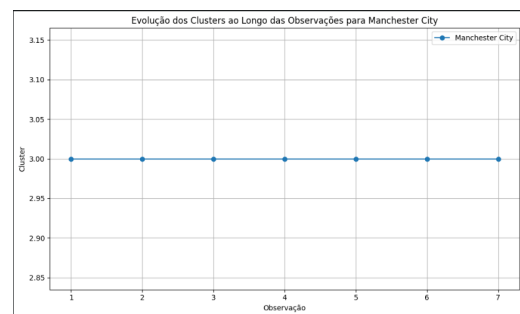


Figure 10. Clusters do M. City

o Manchester United demonstrou uma variação interessante ao longo dos anos analisados, alternando-se entre os Clusters 0 e 3. Essa alternância entre clusters sugere que o Manchester United oscilou entre períodos com um ataque prolífico e uma defesa estável (Cluster 0) e outros com um ataque muito forte e uma defesa resiliente (Cluster 3). Essa capacidade de transitar entre diferentes perfis de desempenho ao longo dos anos reflete a habilidade do Manchester United de se adaptar e responder às condições específicas de cada temporada, buscando maximizar sua eficácia tanto na criação de chances quanto na defesa, em busca de resultados positivos.

Ao longo dos anos analisados, o West Ham apresentou uma notável variação entre os Clusters 0, 1 e 2, nunca sendo atribuído ao Cluster 3. Houve um período de estabilidade significativa no Cluster 1, caracterizado por um equilíbrio moderado entre Expected Goals (xG) e Expected Goals Against (xGA), refletindo em uma acumulação consistente de pontos.

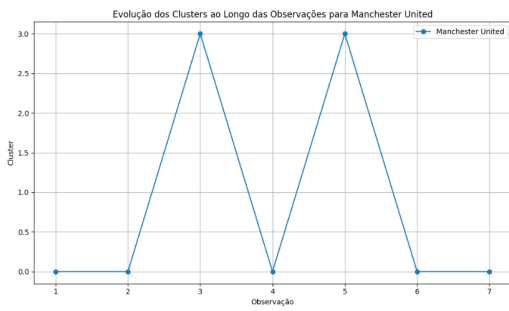


Figure 11. Clusters do M. United

Essa variabilidade sugere uma adaptação do estilo de jogo do clube ao longo das temporadas, explorando diferentes perfis de desempenho para alcançar resultados variados.

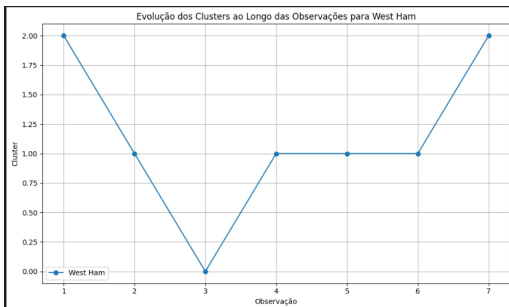


Figure 12. Clusters do West Ham

7.2 Times que mais aparecem em cada cluster

Cluster 0: Times como Tottenham, Manchester United, Aston Villa, Everton e Arsenal mostram equilíbrio entre criação de oportunidades de gol (alta média de xG) e defesa sólida (baixo xGA), resultando em boa média de pontos.

Cluster 1: Crystal Palace, Leicester, West Ham, Burnley e Fulham criam menos oportunidades de gol com uma defesa razoável, resultando em média de pontos mais baixa.

Cluster 2: Burnley, Sheffield United, Fulham, Luton e West Ham têm dificuldades tanto na criação de oportunidades quanto na defesa, refletindo em média de pontos mais baixa.

Cluster 3: Manchester City, Arsenal, Chelsea, Liverpool e Manchester United destacam-se pela alta capacidade de criar chances de gol e defesa sólida, resultando em média de pontos robusta

8 Análise de sentimentos por meio de tweets

O Manchester United foi identificado como o clube que recebeu o maior número de tweets negativos em comparação com outros times analisados. Os clubes mais frequentemente mencionados em tweets, destacando-se por sua presença significativa nas redes sociais, incluem Arsenal, Manchester

```
Times que mais aparecem em cada cluster:
Cluster 0:
['Tottenham', 'Manchester United', 'Aston Villa', 'Everton', 'Arsenal']

Cluster 1:
['Crystal Palace', 'Leicester', 'West Ham', 'Burnley', 'Fulham']

Cluster 2:
['Burnley', 'Sheffield United', 'Fulham', 'Luton', 'West Ham']

Cluster 3:
['Manchester City', 'Arsenal', 'Chelsea', 'Liverpool', 'Manchester United']
```

Figure 13. Clusters

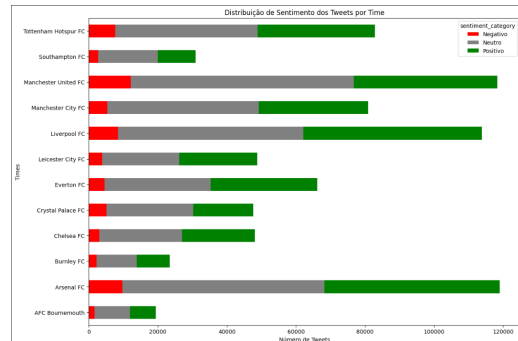


Figure 14. Análise Sentimento

United, Liverpool e Tottenham. Esses times foram consistentemente objeto de discussões e críticas por parte dos usuários, refletindo um alto nível de engajamento e interação nas plataformas digitais.

Além disso, a figura abaixo esboça um gráfico de linhas com a evolução dos tweets de alguns times ao longo do tempo

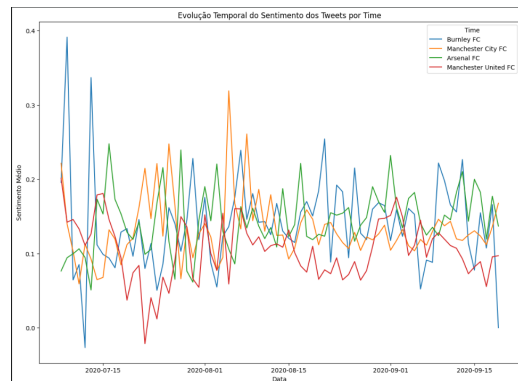


Figure 15. Análise Sentimento ao Longo do Tempo

Ao longo da análise dos tweets ao longo do tempo, foram observados vários picos significativos de atividade. Esses picos indicam momentos de aumento substancial na quantidade de tweets relacionados aos clubes de futebol analisados. Essas flutuações podem refletir eventos específicos, como partidas importantes, transferências de jogadores, eventos extracampo, ou até mesmo momentos de alta emoção durante a temporada. Esses picos são indicativos do impacto dinâmico e da natureza volátil das discussões e interações nas redes sociais em torno desses times de futebol.

8.1 Entendendo comentários negativos do Manchester United

Para uma análise mais profunda dos sentimentos negativos em relação ao Manchester United, foram coletados dados detalhados das partidas ocorridas em 2020. Esses dados fornecem informações específicas sobre o desempenho do clube ao longo da temporada, incluindo resultados, desempenho de jogadores, momentos-chave das partidas, entre outros aspectos relevantes. Essa mineração de dados permite uma investigação mais detalhada e contextualizada dos motivos por trás das interações negativas nas redes sociais relacionadas ao Manchester United durante o período analisado.

As derrotas fora de casa, especialmente contra equipes que não estão no topo da tabela, podem ser criticadas como falta de capacidade de lidar com a pressão ou de garantir pontos cruciais.



```
Quantidade de derrotas em casa: 8
Times que derrotaram o Manchester United em casa:
team_a
Crystal Palace      2
Burnley             1
Tottenham           1
Arsenal             1
Sheffield United    1
Leicester           1
Liverpool           1
Name: count, dtype: int64

Quantidade de derrotas fora de casa: 6
Times que derrotaram o Manchester United fora de casa:
team_h
West Ham            1
Newcastle United    1
Bournemouth         1
Watford             1
Arsenal             1
Liverpool           1
Name: count, dtype: int64
```

Figure 16. Análise Manchester United

O Manchester United enfrentou desafios significativos tanto em seus jogos em casa quanto fora durante a temporada. Em Old Trafford, o time sofreu derrotas notáveis para equipes como o Sheffield e enfrentou rivais tradicionais, destacando a intensidade e competitividade dos confrontos em seu próprio estádio. Fora de casa, o Manchester United também encontrou dificuldades, acumulando derrotas contra seis equipes diferentes, revelando a exigência do calendário e a competitividade nos estádios adversários ao redor da liga.

References

- [1] 2024. *Understat*. Retrieved June 17, 2024 from <https://understat.com/>
[1]