

# Dados na mesa do bar

---

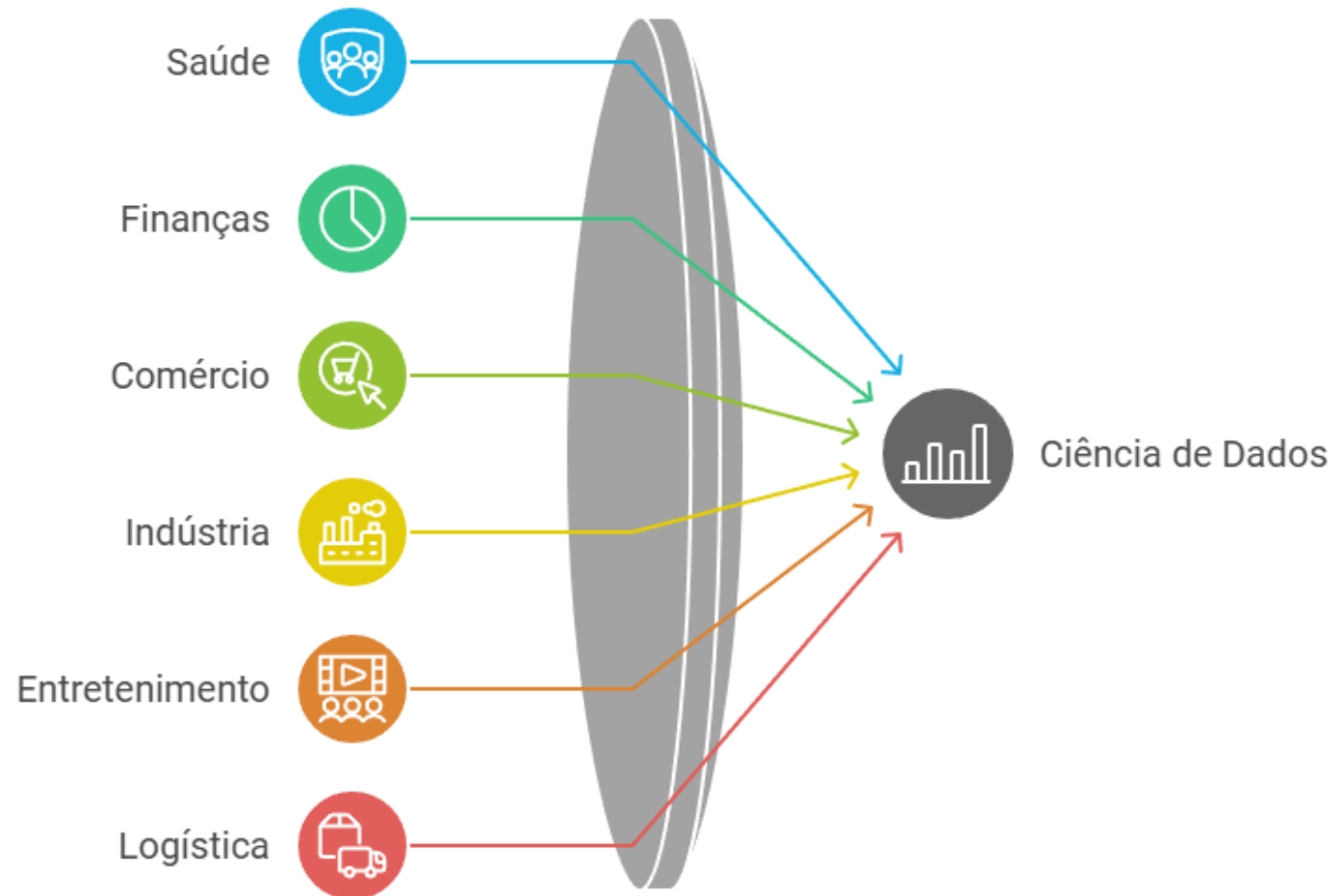
A ciência por trás dos  
palpites



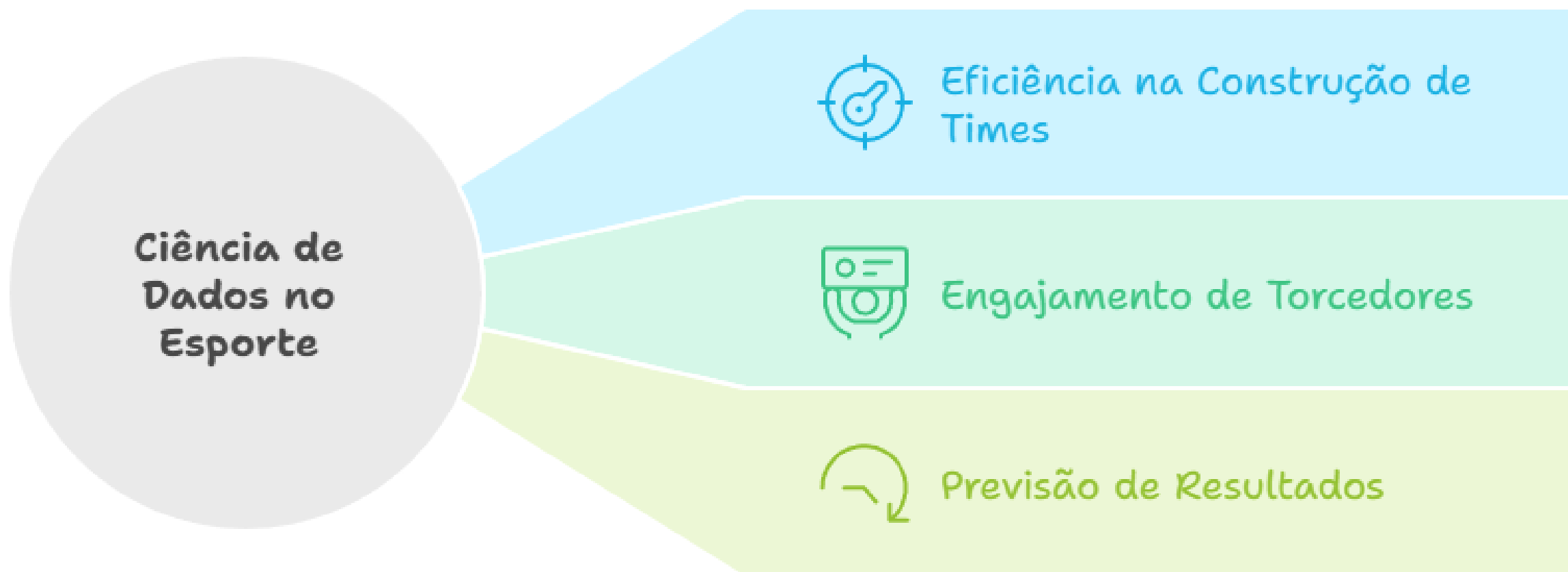
André Luiz França Batista  
@sigaandreluiz



# Ciência de Dados



# Ciência de Dados nos esportes



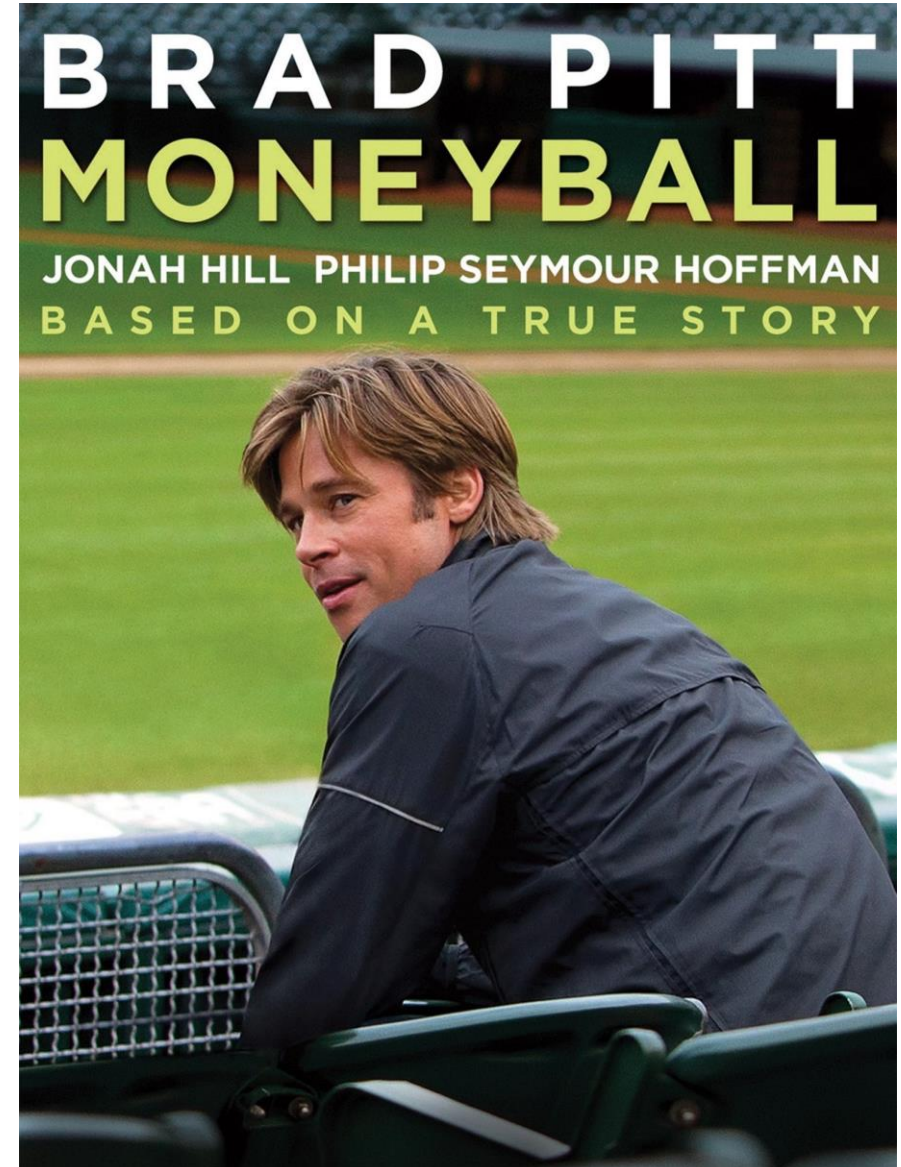
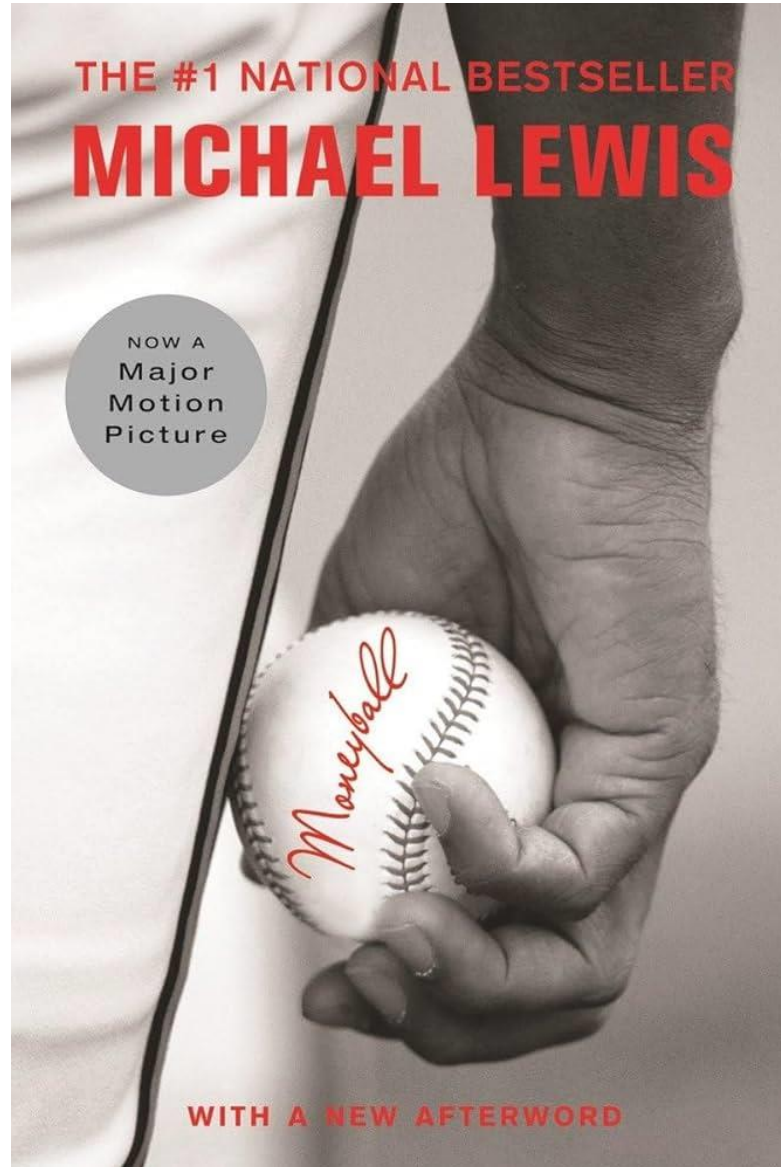


# **Eficiência na formação de equipes**





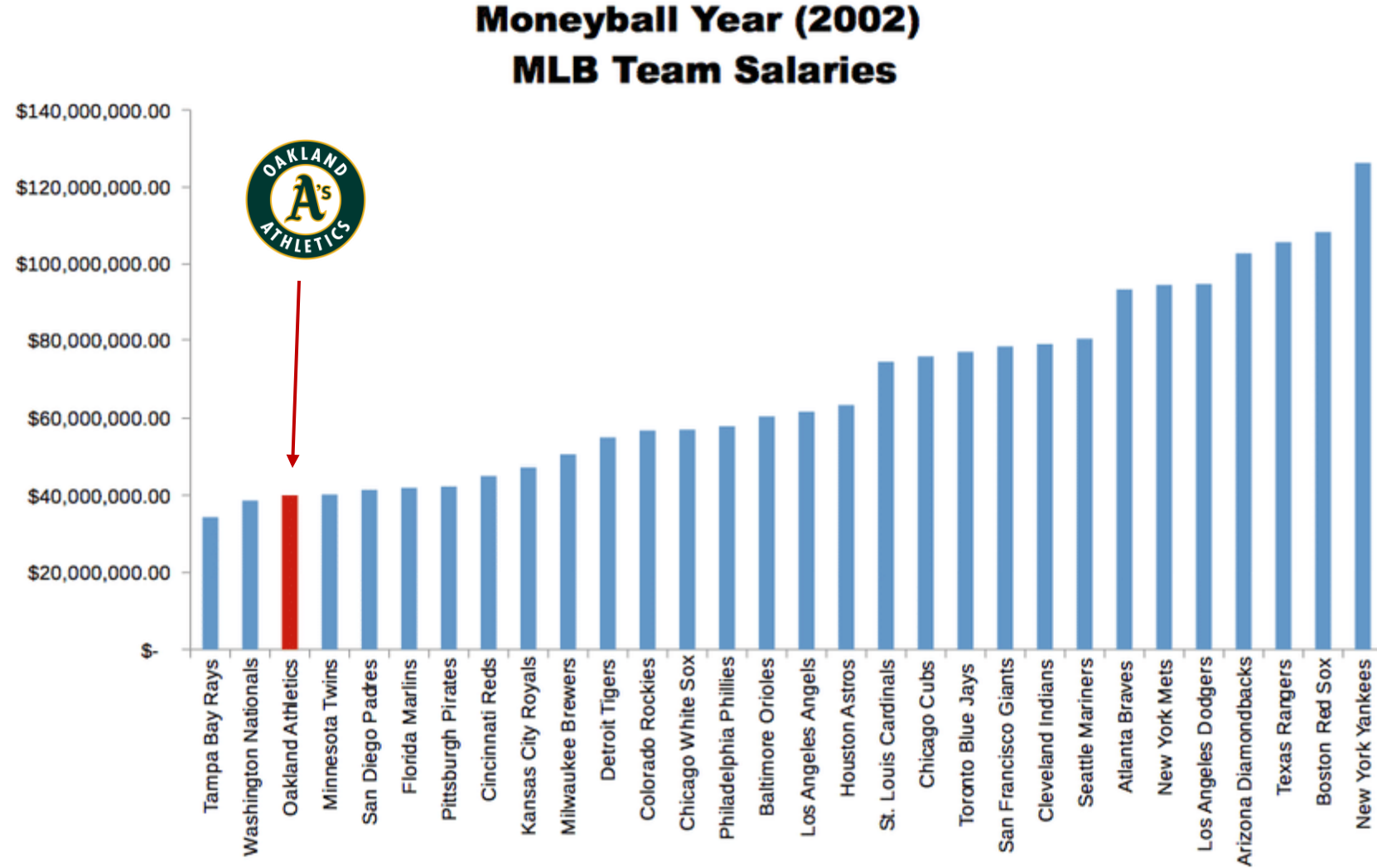




# Problema do Oakland A's

---

- Ano de 2002
- Corte de salários
- Buscar eficiência nos gastos
- Buscar vitórias sem jogadores “estrelas”
- Contrataram um economista: Paul DePodesta



# Resumo da história

---

- Vinte vitórias seguidas (4th na história)
- Não ganhou a World Series
- Fez história e mudou o “rumo” do jogo por meio da análise
- Outros times passaram a usar análise de dados/estatísticas para otimizar os gastos com jogadores









## FOOTBALLER KEVIN DE BRUYNE USES DATA ANALYTICS IN NEW CONTRACT NEGOTIATIONS

Could this new approach become the norm in professional sports?

In negotiating his latest contract, Manchester City star **Kevin De Bruyne**, has done something that's never been seen in professional football, eschewing an agent and hiring data analysts instead. We explore why he took this unconventional approach, and why it might be a glimpse at the future of professional sports contracts.

### Who is Kevin De Bruyne?

Kevin De Bruyne is a world-class footballer, almost in a category of his own. Alongside past and present luminaries Lionel Messi, Zinedine Zidane and Ronaldinho.



**Kevin de Bruyne's** contract negotiations with his club Manchester City has made headlines all over the world in this past week, note **Diario AS**. The Belgian midfielder has taken the reins of his career more firmly and assertively than anyone before, commissioning a data firm to break down with finite detail the value he brings to **Pep Guardiola's** team.



## Kevin De Bruyne and the rise of data and analytics in contract negotiations







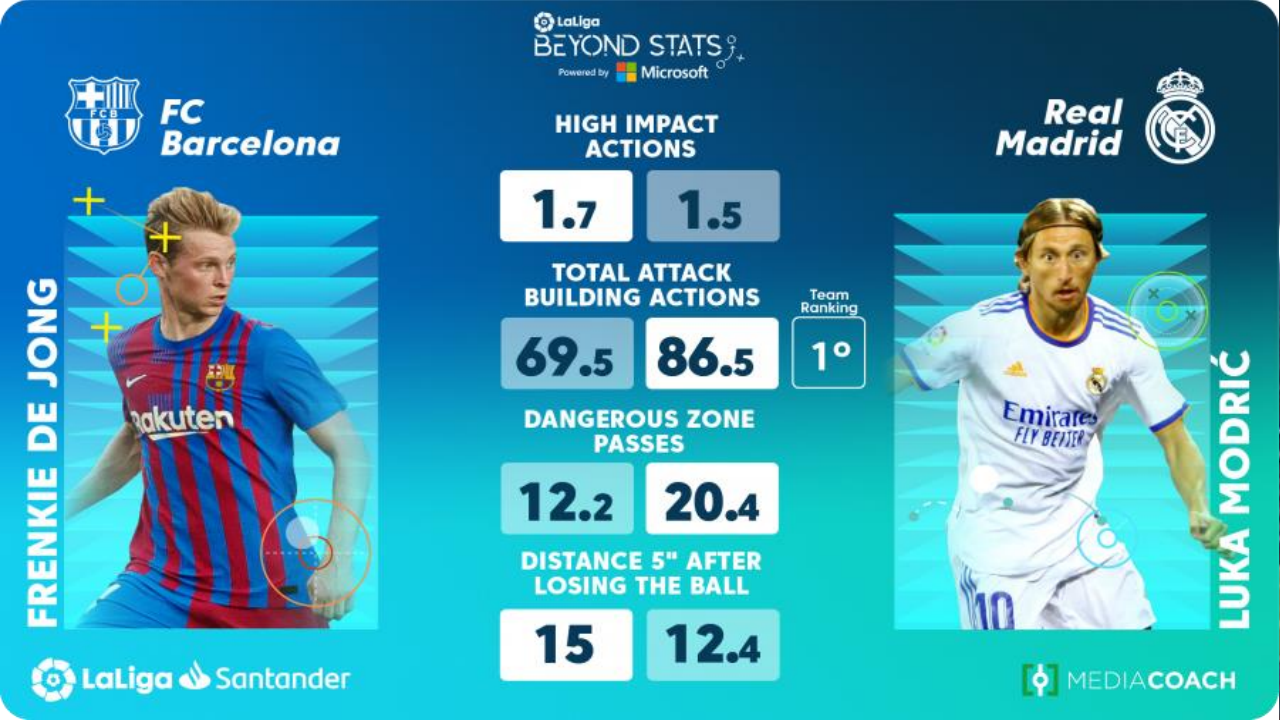






# **Engajamento de torcedores**







# Ciência de Dados e Storytelling

Usar dados para dar aos torcedores uma melhor compreensão do jogo e torná-los mais engajados com o conteúdo.



04/05/2022 • 16:00



Real Madrid

Agg: 6 - 5

**3 - 1**

After extra time



Man City

K. Benzema 95' (Pen.)



R. Mahrez 73'

Rodrygo 90'+1, 90'

Full time

1  
▲ 3.50

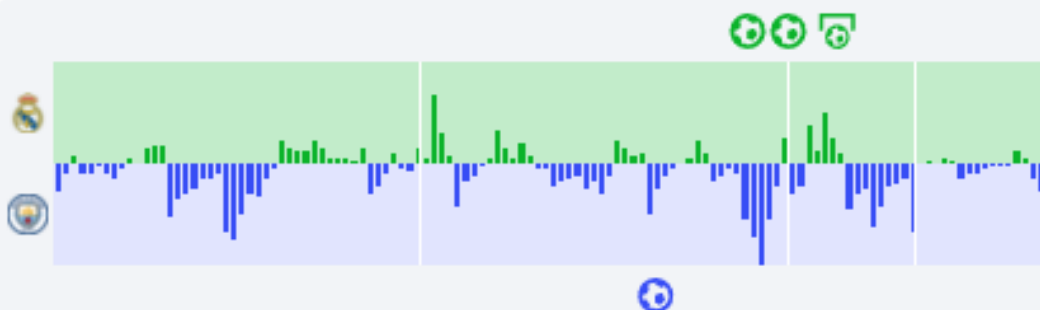
X  
▲ 4.00

2  
▼ 1.91

1XBET

Gamble responsibly 18+

Additional odds



FT 2 - 1



90'+6

F. Valverde Foul



90'+1

**2 - 1** Rodrygo Assist: M. Asensio



90'

**1 - 1** Rodrygo Assist: K. Benzema

Additional time 6'

Out: R. Mahrez In: Fernandinho

85'



84'

E. Militão Foul

Out: G. Jesus In: J. Grealish

78'



75'

In: E. Camavinga Out: L. Modrić



75'

In: M. Asensio Out: Casemiro

Assist: B. Silva R. Mahrez **0 - 1**

73'



Out: K. De Bruyne In: I. Gündoğan

72'



Out: K. Walker In: O. Zinchenko

72'



68'

In: Rodrygo Out: T. Kroos



65'

D. Carvajal Foul

HT 0 - 0



# Previsão de resultados



# **Modelos preditivos**

- **Modelo híbrido de duas etapas.**
- **Modelo sistemas lineares com mínimos quadrados.**

# Modelo híbrido de duas etapas

**Estratégia:** utilizar  
um processo  
sequencial

Etapa 1

Cálculo de Forças e  
Probabilidades Iniciais

Etapa 2

Classificação com Machine  
Learning



# Como damos uma “nota” para cada time?

- Analisamos os resultados de muitos jogos que já aconteceram (quem ganhou, perdeu ou empatou).
- Cada time recebe uma "nota" que representa o quão forte ele é.

Liga	Time	Força ▾
T1	Galatasaray	3.6409976426748503
F1	Paris SG	2.8601829467078397
P1	Sp Lisbon	2.725833149651435
E2	Birmingham	2.699729218833392
D1	Bayern Munich	2.6656646444707763
T1	Fenerbahce	2.659733269111279
SC0	Celtic	2.6579127031928773
G1	Olympiakos	2.6496885263975574
P1	Benfica	2.5226522912118376
E0	Liverpool	2.4188129497777866
SP1	Barcelona	2.311143486317576



# Como estimamos as chances?

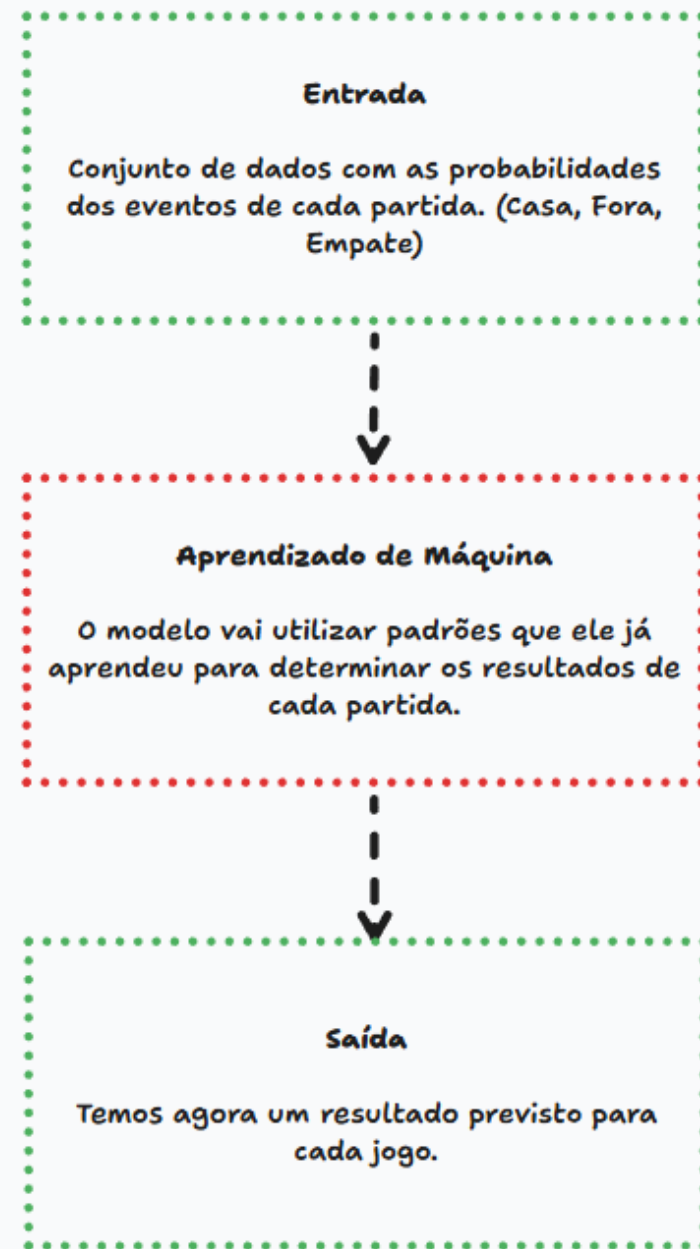
3. Com a "pontuação de força" de cada time em um jogo, calculamos as probabilidades.

Data	Liga	Casa	Fora	ProbCasa	ProbEmpate	ProbFora
17/05/2025	D1	Hoffenheim	Bayern Munich	0.022969596847142	0.1639888442329022	0.8130415589199557
17/05/2025	F1	Paris SG	Auxerre	0.7738756529573465	0.192776117443549	0.0333482295991044
17/05/2025	P1	Porto	Nacional	0.7032305093966871	0.2399247840505284	0.0568447065527846
17/05/2025	SC0	Celtic	St Mirren	0.7391408936322578	0.2167283243305922	0.0441307820371499
18/05/2025	E0	Everton	Southampton	0.7496884645417055	0.2096120330197681	0.0406995024385264
18/05/2025	I1	Parma	Napoli	0.0330024724924837	0.1919226828357928	0.7750748446717234
18/05/2025	N1	AZ Alkmaar	Almere City	0.7157838302865466	0.2319977959466058	0.0522183737668476
18/05/2025	N1	Sparta Rott...	PSV Eindhoven	0.0510792075151339	0.2299616548475817	0.7189591376372843
18/05/2025	T1	Fenerbahce	Eyupspor	0.7074916881357545	0.2372559896013548	0.0552523222628906
18/05/2025	T1	Galatasaray	Kayserispor	0.8285687945667756	0.1520535203468414	0.019377685086383

# O “juiz inteligente”

## Etapa 2

Apenas probabilidades acima de 70% são inseridas no modelo.



# Resultado final

**Legenda:**  
H: Casa ou Empate  
A: Fora ou Empate

Data	Liga	Casa	Fora	ProbCasa	ProbEmpate	ProbFora	Previsão	
17/05/2025	D1	Hoffenheim	Bayern Munich	0.022969596847142	0.1639888442329022	0.8130415589199557	A	●
17/05/2025	F1	Paris SG	Auxerre	0.7738756529573465	0.192776117443549	0.0333482295991044	H	●
17/05/2025	P1	Porto	Nacional	0.7032305093966871	0.2399247840505284	0.0568447065527846	H	●
17/05/2025	SC0	Celtic	St Mirren	0.7391408936322578	0.2167283243305922	0.0441307820371499	H	●
18/05/2025	E0	Everton	Southampton	0.7496884645417055	0.2096120330197681	0.0406995024385264	H	●
18/05/2025	I1	Parma	Napoli	0.0330024724924837	0.1919226828357928	0.7750748446717234	A	●
18/05/2025	N1	AZ Alkmaar	Almere City	0.7157838302865466	0.2319977959466058	0.0522183737668476	H	●
18/05/2025	N1	Sparta Rott...	PSV Eindhoven	0.0510792075151339	0.2299616548475817	0.7189591376372843	A	●
18/05/2025	T1	Fenerbahce	Eyupspor	0.7074916881357545	0.2372559896013548	0.0552523222628906	H	●
18/05/2025	T1	Galatasaray	Kayserispor	0.8285687945667756	0.1520535203468414	0.019377685086383	H	●



# Desempenho do modelo

Liga	País	Casa (H)	Fora (A)	Total
Jupiler Pro League	Bélgica	87.50% (8)	87.50% (8)	87.50% (16)
Bundesliga	Alemanha	93.75% (16)	100.00% (15)	96.77% (31)
Premier League	Inglaterra	100.00% (32)	93.75% (32)	96.88% (64)
Championship	Inglaterra	100.00% (14)	100.00% (14)	100.00% (28)
League One	Inglaterra	100.00% (23)	82.61% (23)	91.30% (46)
FA Cup	Escócia	100.00% (10)	90.00% (10)	95.00% (20)
Ligue 1	França	100.00% (18)	89.47% (19)	94.59% (37)
Ligue 2	França	100.00% (4)	100.00% (4)	100.00% (8)
Super League Grega	Grécia	100.00% (14)	85.71% (14)	92.86% (28)
Serie A	Itália	100.00% (27)	93.10% (29)	96.43% (56)
Serie B	Itália	100.00% (1)	100.00% (1)	100.00% (2)
Eredivisie	Holanda	93.33% (30)	90.00% (30)	91.67% (60)
Primeira Liga	Portugal	100.00% (32)	84.85% (33)	92.31% (65)
Premiership Escocesa	Escócia	100.00% (16)	93.75% (16)	96.88% (32)
Championship Escocesa	Escócia	100.00% (4)	100.00% (4)	100.00% (8)
La Liga	Espanha	91.30% (23)	82.61% (23)	86.96% (46)
Segunda División	Espanha	80.00% (10)	66.67% (9)	73.68% (19)
Süper Lig	Turquia	100.00% (39)	95.24% (42)	97.53% (81)

# Modelo preditivo com sistemas lineares

*Determinado*

$$\begin{cases} x + y = 30 \\ x - y = 10 \end{cases}$$

$$x = 20$$

$$y = 10$$

*Sobredeterminado*

$$\begin{cases} 3a - 2b = 1 & a = ? \\ 2c - 4d = -2 & b = ? \\ 7e - 4f = 3 & c = ? \\ 6g - 3h = 3 & d = ? \\ 1i - 1j = 0 & \dots \\ 4a - 1f = 3 \\ 4b - 3g = 1 \\ 1c - 2h = -1 \\ 2d - 2i = 0 \\ 0e - 1j = -1 \end{cases}$$

# Solução com mínimos quadrados

Resolve um sistema linear da forma:

$$Ax \approx b$$

Ou seja, quando o sistema não tem uma solução exata (por exemplo, em sistemas **sobredeterminados** — mais equações do que variáveis), ele encontra a **solução que minimiza o erro quadrático**:

$$\min_x \|Ax - b\|^2$$



# Metodologia

1. Conversão da tabela de resultados para um sistema linear sobredeterminado;
2. Aplicação do método de mínimos quadrados;
3. Cada time (variável) ganha um score (força);

$$\left\{ \begin{array}{l} 0a - 0b = 0 \\ 2c - 0d = 2 \\ 2e - 0f = 2 \\ 2g - 1h = 1 \\ 2i - 1j = 1 \\ 0k - 0l = 0 \\ 2m - 1n = 1 \\ 1o - 1p = 0 \\ 1q - 1r = 0 \\ 2s - 2t = 0 \\ 3p - 0m = 3 \\ 2t - 0i = 2 \\ 0j - 0a = 0 \\ 2f - 1s = 1 \\ 1e - 1h = 0 \\ 3r - 0g = 3 \\ 1d - 2q = -1 \\ 1b - 2k = -1 \\ 2n - 2o = 0 \\ 2l - 0c = 2 \end{array} \right.$$

Home	Score	Away
<a href="#">São Paulo</a>	0-0	<a href="#">Sport Recife</a>
<a href="#">Juventude</a>	2-0	<a href="#">Vitória</a>
<a href="#">Fortaleza</a>	2-0	<a href="#">Fluminense</a>
<a href="#">Cruzeiro</a>	2-1	<a href="#">Mirassol</a>
<a href="#">Grêmio</a>	2-1	<a href="#">Atlético Mineiro</a>
<a href="#">Palmeiras</a>	0-0	<a href="#">Botafogo (RJ)</a>
<a href="#">Vasco da Gama</a>	2-1	<a href="#">Santos</a>
<a href="#">Bahia</a>	1-1	<a href="#">Corinthians</a>
<a href="#">Flamengo</a>	1-1	<a href="#">Internacional</a>
<a href="#">RB Bragantino</a>	2-2	<a href="#">Ceará</a>
<a href="#">Corinthians</a>	3-0	<a href="#">Vasco da Gama</a>
<a href="#">Ceará</a>	2-0	<a href="#">Grêmio</a>
<a href="#">Atlético Mineiro</a>	0-0	<a href="#">São Paulo</a>
<a href="#">Fluminense</a>	2-1	<a href="#">RB Bragantino</a>
<a href="#">Mirassol</a>	1-1	<a href="#">Fortaleza</a>
<a href="#">Internacional</a>	3-0	<a href="#">Cruzeiro</a>
<a href="#">Vitória</a>	1-2	<a href="#">Flamengo</a>
<a href="#">Sport Recife</a>	1-2	<a href="#">Palmeiras</a>
<a href="#">Santos</a>	2-2	<a href="#">Bahia</a>
<a href="#">Botafogo (RJ)</a>	2-0	<a href="#">Juventude</a>

Data	Hora	Casa	Score		Score	Visitante	Diff	Placar	Status
15/05/2025	14:00	Osasuna	-0.14	vs	0.82	Ath Madrid	-0.96	-	●
15/05/2025	14:00	Vallecano	-0.13	vs	0.25	Betis	-0.38	-	●
15/05/2025	16:30	Espanol	-0.27	vs	1.58	Barcelona	-1.85	-	●
15/05/2025	16:30	Getafe	-0.13	vs	0.68	Ath Bilbao	-0.80	-	●
16/05/2025	15:30	Aston Villa	0.19	vs	0.10	Tottenham	0.09	-	●
16/05/2025	16:15	Chelsea	0.50	vs	-0.31	Man United	0.81	-	●
17/05/2025	08:30	Celtic	2.17	vs	-0.13	St Mirren	2.30	-	●

3. Cada time (variável) ganha um Score (força);
4. Subtrair Score da casa pelo Score do visitante, obtém a diferença *Diff*;
5. Filtrar jogos com *Diff* absoluto acima de 1.4;
  - *Diff* positivo: Casa ou empate.
  - *Diff* negativo: Visitante ou empate.





# Previsão com filtro



































Data	Hora	Casa	Score		Score	Visitante	Diff	Placar	Status
15/05/2025	16:30	Espanol	-0.27	vs	1.58	Barcelona	-1.85	0x2	●
17/05/2025	08:30	Celtic	2.17	vs	-0.13	St Mirren	2.30	1x1	●
17/05/2025	10:30	Hoffenheim	-0.57	vs	1.82	Bayern Munich	-2.38	0x4	●
17/05/2025	15:45	Genoa	-0.38	vs	1.10	Atalanta	-1.48	2x3	●
17/05/2025	16:00	Paris SG	1.57	vs	-0.07	Auxerre	1.65	3x1	●
18/05/2025	08:00	Everton	-0.11	vs	-1.52	Southampton	1.41	2x0	●
18/05/2025	10:30	Greuther Furth	-0.46	vs	1.01	Hamburg	-1.47	3x2	●
18/05/2025	13:00	Galatasaray	1.51	vs	-0.27	Kayserispor	1.78	3x0	●
25/05/2025	10:00	Atalanta	1.10	vs	-0.45	Parma	1.55	-	●
25/05/2025	10:00	Milan	0.49	vs	-0.97	Monza	1.46	-	●
25/05/2025	12:00	Bournemouth	0.33	vs	-1.22	Leicester	1.55	-	●
25/05/2025	12:00	Southampton	-1.52	vs	0.89	Arsenal	-2.41	-	●



# Desempenho 2024/2025

## Legenda

-  Excelente ( $\geq 80\%$ )
-  Bom (65% a 79%)
-  Regular (50% a 64%)
-  Abaixo do ideal ( $< 50\%$ )

 País	 Liga	 Casa (Jogos)	 Fora (Jogos)	 Total (Jogos)
 Inglaterra	Premier League	 85.71% (35)	 90.32% (31)	 <b>87.88%</b> (66)
 Inglaterra	EFL Championship	 100.00% (26) 	 65.62% (32)	 <b>81.03%</b> (58)
 Alemanha	Bundesliga	 88.24% (34)	 81.48% (27)	 <b>85.25%</b> (61)
 Alemanha	2. Bundesliga	 84.21% (19)	 66.67% (27)	 <b>73.91%</b> (46)
 Itália	Serie A	 96.30% (27)	 80.65% (31)	 <b>87.93%</b> (58)
 Itália	Serie B	 75.00% (4)	 70.00% (10)	 <b>71.43%</b> (14)
 França	Ligue 1	 94.12% (34)	 71.79% (39)	 <b>82.19%</b> (73)
 França	Ligue 2	 75.00% (12)	 64.71% (17)	 <b>68.97%</b> (29)
 Espanha	La Liga	 92.00% (25)	 78.57% (28)	 <b>84.91%</b> (53)
 Espanha	La Liga 2	 62.50% (8)	 44.44% (9)	 <b>52.94%</b> (17)

# Recap



Ciência de dados  
no esporte



Formação de  
equipes



Engajamento de  
torcedores

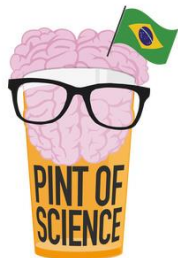


Previsão de  
resultados



**Obrigado!**





# Dados na mesa do bar

A ciência por trás dos  
palpites



André Luiz França Batista  
@sigaandreluiz



Data	Hora	Casa	Score		Score	Visitante	Diff	Placar	Status
19/05/2025	15:00	Racing	0.30	vs	-0.55	Miramar Misiones	0.86	-	●
19/05/2025	16:00	Brighton	0.04	vs	1.21	Liverpool	-1.17	-	●
19/05/2025	16:00	Unión Española	-0.62	vs	-1.15	Everton	0.53	-	●
19/05/2025	19:00	Argentinos Jrs	1.05	vs	0.22	San Lorenzo	0.83	-	●
19/05/2025	19:00	Boston River	-0.04	vs	-0.46	Cerro	0.43	-	●
19/05/2025	19:00	Colo-Colo	0.28	vs	-0.81	Ñublense	1.09	-	●
19/05/2025	19:00	Volta Redonda	-0.57	vs	-0.54	Amazonas	-0.03	-	●
19/05/2025	19:30	Binacional	-0.40	vs	0.00	Cienciano	-0.40	-	●
19/05/2025	21:00	LDU de Quito	0.73	vs	0.01	Libertad	0.72	-	●
19/05/2025	21:30	Boca Juniors	1.01	vs	0.57	Independiente	0.44	-	●
19/05/2025	21:30	La Serena	-0.75	vs	-1.12	Iquique	0.38	-	●
19/05/2025	21:35	Coritiba	0.51	vs	-0.25	América (MG)	0.75	-	●
19/05/2025	22:00	Alianza Lima	0.40	vs	-1.07	Alianza Univ	1.47	-	●
20/05/2025	16:00	Crystal Palace	-0.08	vs	-0.35	Wolves	0.27	-	●
20/05/2025	16:00	Man City	0.62	vs	0.33	Bournemouth	0.29	-	●