

Pontuações do Brasileirão Série A Série Histórica - de 2003 a 2022

Relatório Análise e Projeção a partir da Pontuação

Gabriel do Nascimento Gomes

[Pontuações Notáveis Brasileirão A - Série Histórica.xlsx](#)

• Introdução

(1.0)

O conceito inicial dessa análise de série histórica era apenas projetar através de métodos de análise de dados quais seriam os “números mágicos” (pontuação/aproveitamento que garantiria estar numa determinada faixa da tabela) de cada faixa importante da tabela, que seriam os chamados “G4, G6, G8, e Z4”. Assim como a UFMG usa modelos estatísticos e probabilísticos para calcular as probabilidades de pontuação por faixa todo ano ([Série A – Probabilidades no Futebol](#)).

Analisando a dispersão dos dados, foram encontrados resultados que apresentam a possibilidade da distância entre os times estarem aumentando, pelas próprias projeções de Campeão, Z4, G4, etc. Pois a distância entre os blocos* estaria aumentando independente dos times, uma vez que a análise foi feita a partir da pontuação, não considerando que time havia feito tal pontuação.

• Explicações de Leitura e do Procedimento

(1.1)

- A Base do Cálculo é feita a partir do **APROVEITAMENTO**. 1.11
- Os dados utilizados são de **2003 a 2022**. 1.12
- Onde tem **observação***, é porque é importante (como 2020 ter sido um campeonato atípico pela pandemia). 1.13
- Considerar **número mágico** um número *mirando acima ou abaixo do Desvio Padrão o suficiente para garantir o time dentro da determinada faixa*. 1.14
- O aproveitamento de **2003~2005** não destoa apesar de serem 2 campeonatos de 24 times e mais 1 de 22 times em 2005, sendo a partir de 2006, 20 times, e foram mantidos apenas para as análises de **PROJEÇÃO**. Observe que número de rodadas e pontuação não entrarão na média (inclusive pontuação foi calculado a partir do Aproveitamento). 1.15

• Dados

(1.2)

Pontuações Notáveis Série Histórica.																@EngenHumanas			Tabela 1.2		
Y	Campeão	Rd	pts	pts por rodada	Aproveitamento	4º	pts por rodada	Aproveitamento	6º	pts por rodada	Aproveitamento	8º	pts por rodada	Aproveitamento	16º	pts por rodada	Aproveitamento	ZR	pts por rodada	Aproveitamento	
2003*	Cruzeiro*	44	94	2,14	71,2%	74	1,61	53,6%	72	1,57	52,2%	66	1,43	47,8%	50	1,09	36,2%	50	1,09	36,2%	
2004*	Santos* !	46	89	1,93	64,5%	79	1,72	57,2%	72	1,57	52,2%	67	1,46	48,6%	51	1,11	37,0%	50	1,09	36,2%	
2005*	Corinthians* !	42	81	1,93	64,3%	70	1,67	55,6%	61	1,45	48,4%	60	1,43	47,6%	51	1,21	40,5%	49	1,17	38,9%	
2006	São Paulo	36	74	2,06	68,5%	64	1,68	56,1%	59	1,55	51,8%	55	1,45	48,2%	44	1,16	38,6%	39	1,03	34,2%	
2007	São Paulo	34	73	2,15	71,6%	61	1,61	53,5%	58	1,53	50,9%	55	1,45	48,2%	45	1,18	39,5%	44	1,16	38,6%	
2008	São Paulo !	38	75	1,97	65,8%	65	1,71	57,0%	54	1,42	47,4%	53	1,39	46,5%	44	1,16	38,6%	44	1,16	38,6%	
2009	Flamengo !	38	67	1,76	58,8%	62	1,63	54,4%	57	1,50	50,0%	55	1,45	48,2%	46	1,21	40,4%	45	1,18	39,5%	
2010	Fluminense !	38	71	1,87	62,3%	63	1,66	55,3%	59	1,55	51,8%	56	1,47	49,1%	42	1,11	36,8%	42	1,11	36,8%	
2011	Corinthians !	38	71	1,87	62,3%	61	1,61	53,5%	59	1,55	51,8%	57	1,50	50,0%	43	1,13	37,7%	41	1,08	36,0%	
2012	Fluminense	35	76	2,17	72,4%	66	1,74	57,9%	57	1,50	50,0%	53	1,39	46,5%	45	1,18	39,5%	41	1,08	36,0%	
2013	Cruzeiro	34	74	2,18	72,5%	61	1,61	53,5%	59	1,55	51,8%	57	1,50	50,0%	45	1,18	39,5%	44	1,16	38,6%	
2014	Cruzeiro	36	76	2,11	70,4%	69	1,82	60,5%	61	1,61	53,5%	54	1,42	47,4%	40	1,05	35,1%	38	1,00	33,3%	
2015	Corinthians	35	77	2,20	73,3%	62	1,63	54,4%	59	1,55	51,8%	55	1,45	48,2%	43	1,13	37,7%	42	1,11	36,8%	
2016	Palmeiras +	37	77	2,08	69,4%	62	1,63	54,4%	57	1,50	50,0%	53	1,39	46,5%	45	1,18	39,5%	43	1,13	37,7%	
2017	Corinthians	35	71	2,03	67,6%	62	1,63	54,4%	56	1,47	49,1%	54	1,42	47,4%	43	1,13	37,7%	43	1,13	37,7%	
2018	Palmeiras +	37	77	2,08	69,4%	66	1,74	57,9%	59	1,55	51,8%	53	1,39	46,5%	43	1,13	37,7%	42	1,11	36,8%	
2019	Flamengo +	34	81	2,38	79,4%	65	1,71	57,0%	63	1,66	55,3%	56	1,47	49,1%	39	1,03	34,2%	36	0,95	31,6%	
2020*	Flamengo* !	38	71	1,87	62,3%	66	1,74	57,9%	59	1,55	51,8%	54	1,42	47,4%	41	1,08	36,0%	41	1,08	36,0%	
2021	Atlético-MG*	36	81	2,25	75,0%	58	1,53	50,9%	56	1,47	49,1%	53	1,39	46,5%	46	1,21	40,4%	43	1,13	37,7%	
2022	Palmeiras	35	77	2,20	73,3%	65	1,71	57,0%	58	1,53	50,9%	55	1,45	48,2%	41	1,08	36,0%	37	0,97	32,5%	
Média	Média	36	75	2,06	69%	63	1,67	55,6%	58	1,53	51,1%	55	1,44	47,9%	43	1,14	37,9%	41	1,09	36,5%	
DP	Desvio Padrão	1,5	4	0,15	7%	3	0,07	4%	2	0,05	3%	1,3	0,03	2%	2	0,05	5%	3	0,06	6%	
V	Variação	2,2	13	0,02	1%	7	0,004	0,3%	4	0,003	0,2%	2	0,0011	0,1%	4	0,0029	0,3%	6	0,004	0,4%	

Teremos como base, a [Rodada em que o time Campeão](#) (pois muitos times relaxam e perdem desempenho após serem campeões) para o time que foi campeão, às demais posições da tabela, usaremos sempre a última rodada como base.

• Análise

(1.3)

Analisando apenas a média e desvio padrão, obtemos essa nova tabela com cores que indicam o quão acima ou abaixo do DP está a pontuação daquele ano:

Tabela 1.31

Pontuações Notáveis Série Histórica:																			@ EngenHumanas		
Y	Campeão	Rd	pts	pts por rodada	Aprovei- tamento	4º	pts por rodada	Aprovei- tamento	6º	pts por rodada	Aprovei- tamento	8º	pts por rodada	Aprovei- tamento	16º	pts por rodada	Aprovei- tamento	ZR	pts por rodada	Aprovei- tamento	
2003*	Cruzeiro*	44	94	2,14	71,2%	74	1,61	53,6%	72	1,57	52,2%	66	1,43	47,8%	50	1,09	36,2%	50	1,09	36,2%	
2004*	Santos* !	46	89	1,93	64,5%	79	1,72	57,2%	72	1,57	52,2%	67	1,46	48,6%	51	1,11	37,0%	50	1,09	36,2%	
2005*	Corinthians* !	42	81	1,93	64,3%	70	1,67	55,6%	61	1,45	48,4%	60	1,43	47,6%	51	1,21	40,5%	49	1,17	38,9%	
2006	São Paulo	36	74	2,06	68,5%	64	1,68	56,1%	59	1,55	51,8%	55	1,45	48,2%	44	1,16	38,6%	39	1,03	34,2%	
2007	São Paulo	34	73	2,15	71,6%	61	1,61	53,5%	58	1,53	50,9%	55	1,45	48,2%	45	1,18	39,5%	44	1,16	38,6%	
2008	São Paulo !	38	75	1,97	65,8%	65	1,71	57,0%	54	1,42	47,4%	53	1,39	46,5%	44	1,16	38,6%	44	1,16	38,6%	
2009	Flamengo !	38	67	1,76	58,8%	62	1,63	54,4%	57	1,50	50,0%	55	1,45	48,2%	46	1,21	40,4%	45	1,18	39,5%	
2010	Fluminense !	38	71	1,87	62,3%	63	1,66	55,3%	59	1,55	51,8%	56	1,47	49,1%	42	1,11	36,8%	42	1,11	36,8%	
2011	Corinthians !	38	71	1,87	62,3%	61	1,61	53,5%	59	1,55	51,8%	57	1,50	50,0%	43	1,13	37,7%	41	1,08	36,0%	
2012	Fluminense	35	76	2,17	72,4%	66	1,74	57,9%	57	1,50	50,0%	53	1,39	46,5%	45	1,18	39,5%	41	1,08	36,0%	
2013	Cruzeiro	34	74	2,18	72,5%	61	1,61	53,5%	59	1,55	51,8%	57	1,50	50,0%	45	1,18	39,5%	44	1,16	38,6%	
2014	Cruzeiro	36	76	2,11	70,4%	69	1,82	60,5%	61	1,61	53,5%	54	1,42	47,4%	40	1,05	35,1%	38	1,00	33,3%	
2015	Corinthians	35	77	2,20	73,3%	62	1,63	54,4%	59	1,55	51,8%	55	1,45	48,2%	43	1,13	37,7%	42	1,11	36,8%	
2016	Palmeiras +	37	77	2,08	69,4%	62	1,63	54,4%	57	1,50	50,0%	53	1,39	46,5%	45	1,18	39,5%	43	1,13	37,7%	
2017	Corinthians	35	71	2,03	67,6%	62	1,63	54,4%	56	1,47	49,1%	54	1,42	47,4%	43	1,13	37,7%	43	1,13	37,7%	
2018	Palmeiras +	37	77	2,08	69,4%	66	1,74	57,9%	59	1,55	51,8%	53	1,39	46,5%	43	1,13	37,7%	42	1,11	36,8%	
2019	Flamengo +	34	81	2,38	79,4%	65	1,71	57,0%	63	1,66	55,3%	56	1,47	49,1%	39	1,03	34,2%	36	0,95	31,6%	
2020*	Flamengo* !	38	71	1,87	62,3%	66	1,74	57,9%	59	1,55	51,8%	54	1,42	47,4%	41	1,08	36,0%	41	1,08	36,0%	
2021	Atlético-MG* +	36	81	2,25	75,0%	58	1,53	50,9%	56	1,47	49,1%	53	1,39	46,5%	46	1,21	40,4%	43	1,13	37,7%	
2022	Palmeiras	35	77	2,20	73,3%	65	1,71	57,0%	58	1,53	50,9%	55	1,45	48,2%	41	1,08	36,0%	37	0,97	32,5%	
Mé din	Média	36	75	2,06	69%	63	1,67	55,6%	58	1,53	51,1%	55	1,44	47,9%	43	1,14	37,9%	41	1,09	36,5%	
DP	Desvio Padrão	1,5	4	0,15	7%	3	0,07	4%	2	0,05	3%	1,3	0,03	2%	2	0,05	5%	3	0,06	6%	
V	Variação	2,2	13	0,02	1%	7	0,004	0,3%	4	0,003	0,2%	2	0,001	0,1%	4	0,003	0,3%	6	0,004	0,4%	

legenda:

- Observação
- ! Campeão na Última Rodada
- + Aproveitamento ainda aumenta, mesmo que momentaneamente após o título
- Melhor
- Acima do DP (ou mt perto)
- Dentro do DP, ACIMA
- Dentro do DP, Perto do Meio
- Dentro da DP, ABAIXO
- Abaixo do DP (ou perigosamente perto)
- Pior

Células Modificáveis

Fontes:

[Rodada Campeão:](#)

[Tabelas de Pontuação:](#)

[Probabilidade UFMG:](#) (Comparar)

Podemos a partir dessa nova tabela diagnosticar rapidamente que existe uma diferença em quantidade de dados acima e abaixo do DP antes e depois de 2012, no caso do Campeão, é mais um indicativo do aumento da diferença dos times. Analisar o atual momento além do todo é importante para qualquer trabalho de projeção. Além de ajudar a determinar para onde a média está se movendo.

Tabela 1.32

Média	Média	36	75	2,06	69%	63	1,67	55,6%	58	1,53	51,1%	55	1,44	47,9%	43	1,14	37,9%	41	1,09	36,5%
DP	Desvio Padrão	1,5	4	0,15	7%	3	0,07	4%	2	0,05	3%	1,3	0,03	2%	2	0,05	5%	3	0,06	6%
V	Variação	2,2	13	0,02	1%	7	0,004	0,3%	4	0,003	0,2%	2	0,001	0,1%	4	0,003	0,3%	6	0,004	0,4%
Média	2003~11	37	73	1,96	65%	63	1,65	55,1%	58	1,52	50,7%	55	1,45	48,3%	44	1,15	38,4%	42	1,12	37,2%
DP	Desvio Padrão	1,5	3	0,12	6%	1	0,04	3%	2	0,05	3%	1,2	0,03	2%	1	0,04	4%	2	0,05	4%
V	Variação	2,3	7	0,01	0,7%	2	0,002	0,1%	3	0,002	0,2%	1	8E-04	0,1%	2	0,002	0,2%	4	0,002	0,2%
Média	2012~22	36	76	2,14	71%	64	1,68	56,0%	59	1,54	51,4%	54	1,43	47,6%	43	1,13	37,6%	41	1,08	35,9%
DP	Desvio Padrão	1,2	3	0,13	6%	3	0,08	5%	2	0,05	3%	1,3	0,03	2%	2	0,06	5%	3	0,07	6%
V	Variação	1,5	10	0,02	0,7%	9	0,006	0,4%	4	0,003	0,2%	2	0,001	0,1%	5	0,003	0,3%	7	0,005	0,4%

Para projetar o *número mágico* foi feito o cálculo de média dos “pontos por rodada”.

P_{mt} → Pontos por Rodada da Média Total

D_{mt} → Desvio Padrão da Média Total

P_{12} → Pontos por Rodada da Média 2012~2022

D_{12} → Desvio Padrão da Média 2012~2022

$$\frac{(P_{mt} + D_{mt}) + (P_{12} + D_{12})}{2}$$

Fórmula 1.33

tabela 1.34

N° Mágico - 2023 - 38Rd			
	Pts	Pts/Rd	Apr
Campeão	85	2,24	75%
G4	66	1,75	58%
G6	60	1,59	53%
G8	56	1,47	49%
16°	46	1,19	40%
ZR	38	1,02	34%

Utilizando esse método relativamente simples de projeção, podemos comparar com o modelo estatístico/probabilístico da UFMG e quão próximo estão esses resultados dos resultados da UFMG.

Tabela 1.35

85	99.999
86	100.000

CAMPEONATO BRASILEIRO SÉRIE A 2023

PONTUAÇÃO NECESSÁRIA PARA SER CAMPEÃO

Tabela 1.36

65	99.999
66	100.000

CAMPEONATO BRASILEIRO SÉRIE A 2023

PONTUAÇÃO NECESSÁRIA PARA SER CLASSIFICADO PARA A LIBERTA

Tabela 1.37

47	0.159
46	0.805
45	3.046
38	93.526
37	97.456
36	99.145
35	99.754
34	99.939
33	99.987
32	99.997
31	100.000

CAMPEONATO BRASILEIRO SÉRIE A 2023
PONTUAÇÃO NECESSÁRIA PARA SER REBAIXADO

• Comparação

(1.4)

Comparando os resultados obtidos com os da UFMG, obtemos o mesmo aproveitamento necessário para ser campeão e G4. A diferença se dá pois 85 pontos é arredondado para baixo, apenas arredondando o valor de $\approx 85,2$ para cima que teríamos o 100% de probabilidade ao invés de “99,999998%” ($\approx 100\%$). O mesmo ocorre para a pontuação necessária para o G4, só que neste caso, o arredondamento foi para cima e os resultados obtidos foram os mesmos.

Para o Z4 houve uma distorção na leitura, uma vez que a pontuação de 31 pontos foi a maior pontuação que um último colocado já obteve, assim probabilisticamente garantindo o rebaixamento com “100%” de precisão (vale lembrar que é um modelo probabilístico). O valor mirado nesse cálculo foi o de garantir ser 17º para baixo, o que probabilisticamente tem mais chance de se salvar do Z4, mas mesmo assim pouca com apenas 38 pontos (arredondado para baixo) tem uma probabilidade de 93,53% de chance de cair. Ao mesmo tempo que 46 pontos não garante se salvar do Rebaixamento mas a probabilidade de ser rebaixado com 46 pontos é de apenas 0,8%.

-Obs¹: Apenas uma vez um time foi rebaixado com 45 pontos, em 2010, dada a **tendência** de queda isso não deve se repetir, então ter 46 pontos como “número mágico” faz todo sentido estatístico.

-Obs²: Apenas uma vez um time se salvou do rebaixamento com 39 pontos, em 2019, justamente o ano da maior distância entre o bloco de cima e debaixo da tabela (veremos abaixo na seção X.X). Ou seja, mesmo assim nenhum time jamais se salvou com 38 pontos. Sendo a **tendência** de queda na pontuação verdadeira, será questão de tempo até alguém se salvar com 38 pontos, ou seja, faz sentido a probabilidade ser menor bem menor e a discrepância entre os resultados serem maiores também.

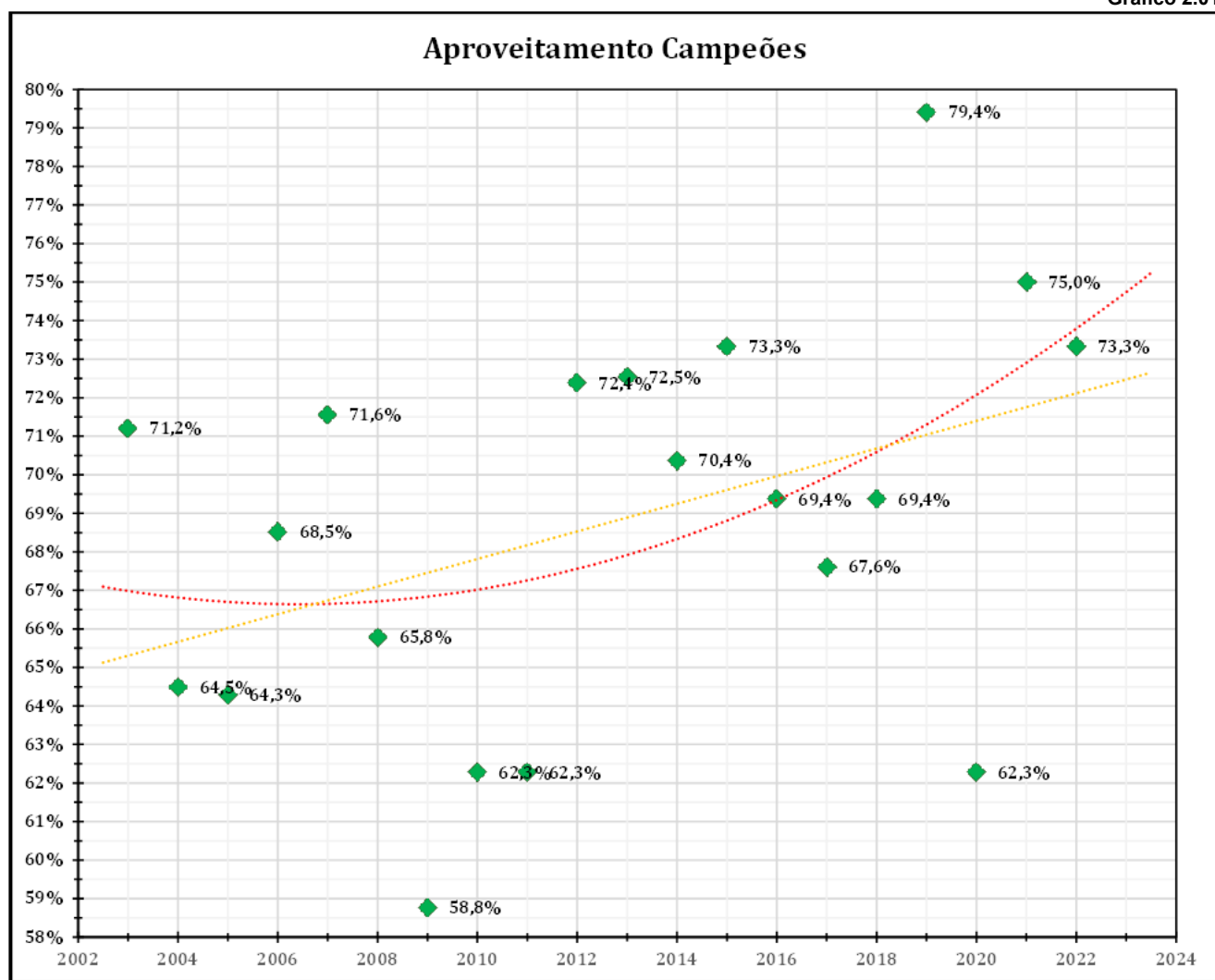
-Obs³: A UFMG não tem pontuação para G6 e G8. Apenas região de Sul-Americana que parte do G10 (55 pts para 100% inclusive).

● Análise

(2.0)

Para observar se esse indício é verdadeiro é necessário analisar os dados e a abordagem de análise realizada foi a do gráfico de dispersão, pois nele podemos ver buscar correlações entre os eixos X e Y (tempo e aproveitamento), e ver realmente se através do tempo (eixo X) a **tendência** da pontuação é subir (aproveitamento, Eixo Y).

Gráfico 2.01

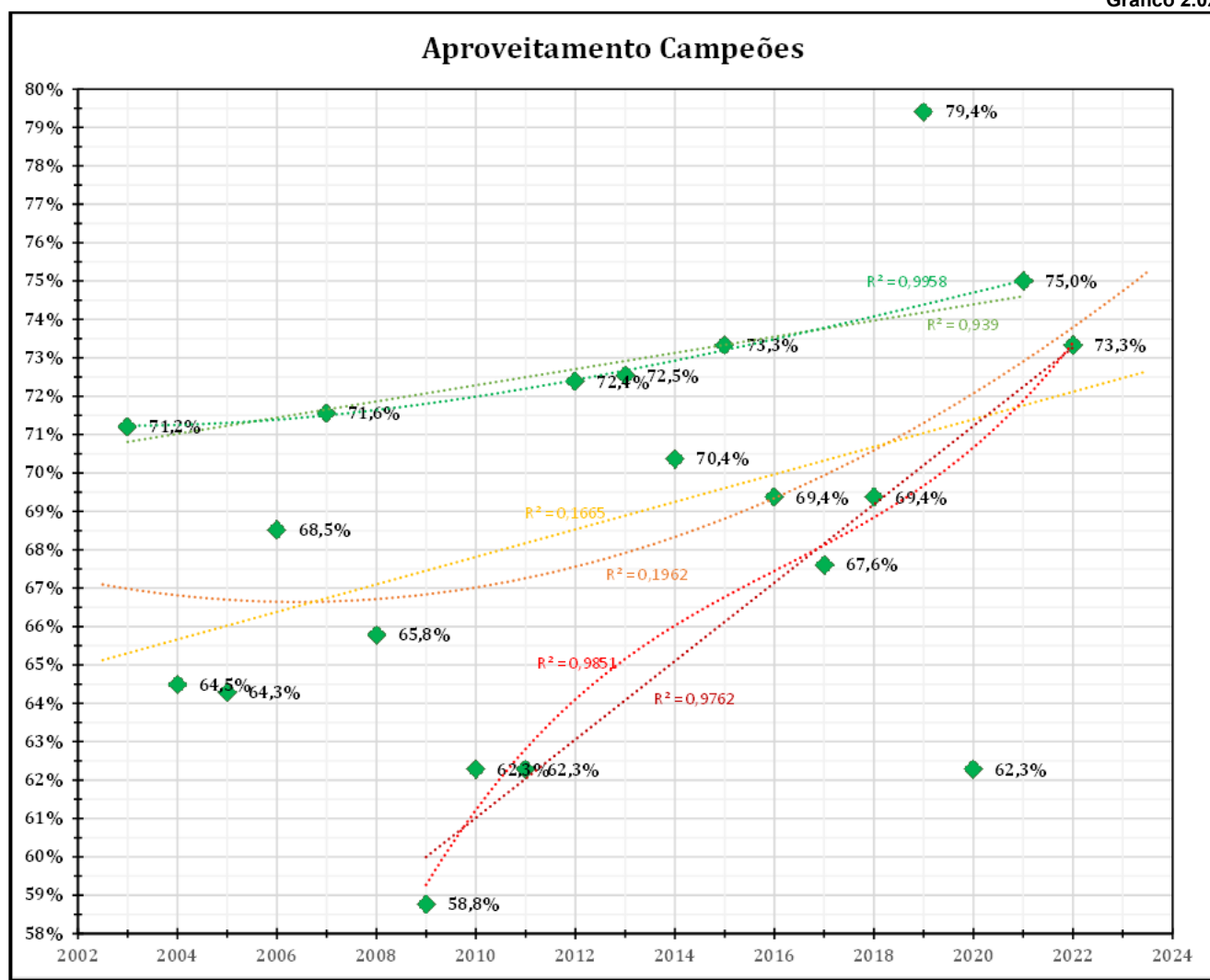


Pelos métodos de ajuste linear e ajuste polinomial de grau 2, os 2 indicam uma **tendência** de crescimento, porém o índice de *correlação* (R^2) foi baixo, $\approx 16,65\%$ e $\approx 19,62\%$ respectivamente. Porém, um índice de *correlação* não significa necessariamente que a **tendência** e correlações não são verdadeiros.

Analisando o gráfico, podemos perceber que a distribuição do setor 2012~18 por exemplo, é mais concentrada (Discrepância Absoluta Máxima de 5,7%) em relação ao setor 2003~11 (Discrepância Absoluta Máxima de 12,8%).

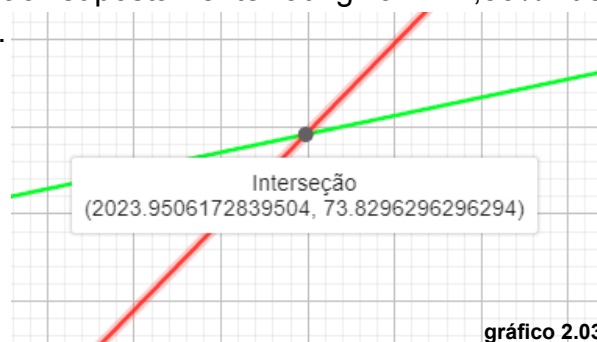
Sendo 2020 um ano atípico pela pandemia, podemos analisar a **tendência** dos “**dados inferiores**” e dos “**dados superiores**” que seriam os dados da parte debaixo e de cima do gráfico respectivamente.

Observe a seguir que agora a *correlação* (R^2) é altíssima desses dados, de $\approx 93,90\%$ a $\approx 99,58\%$. E por uma conclusão lógica, se a parte debaixo, do meio e de cima estão crescendo, e o total também (mesmo com uma “baixa *correlação*”) podemos concluir que há uma **tendência** de crescimento.



Inclusive, no caso do Campeão, há um “afunilamento” dos dados na projeção uma vez que os dados inferiores do gráfico “crescem mais rápido” (maior ângulo de inclinação em relação ao eixo X) que os dados superiores do gráfico. A projeção da interseção dos ajustes lineares acontece em ≈ 2024 , onde o Campeão supostamente atingiria $\approx 74,83\%$ de aproveitamento, o que em pontos significa ≈ 84 pts.

Meta essa atingida apenas recentemente, apenas em 2019 e 2021, por Flamengo (90 pts) e o Atlético-MG (84 pts), sendo esse, mais um indicativo da diferença dos times estarem aumentando.



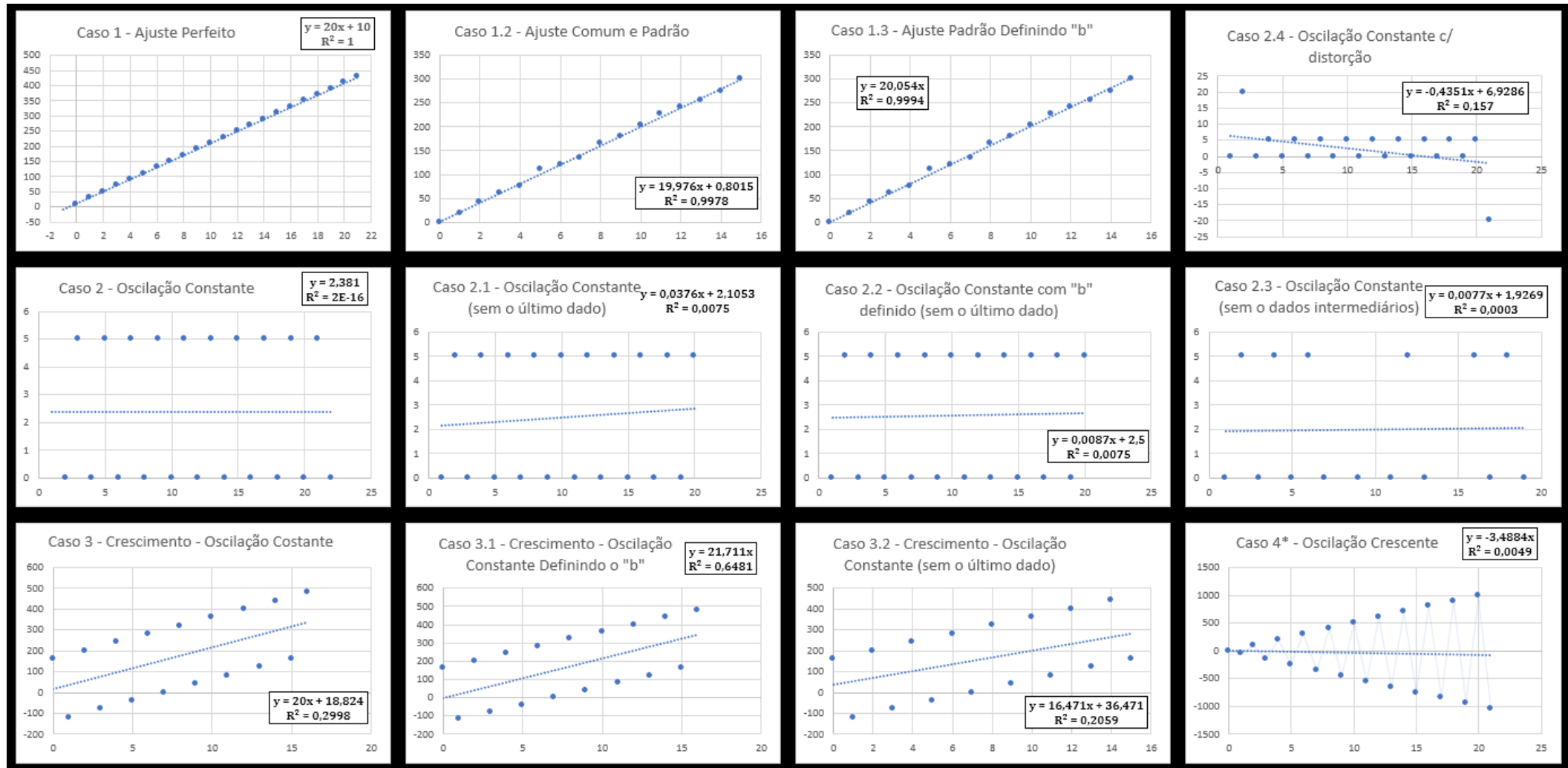
A seguir, veremos alguns tipos de distribuição em gráficos de dispersão e sua suposta *correlação* seguidos de uma explicação, uma vez que como demonstrado no caso anterior, uma baixa “*correlação*” (R^2) não significa necessariamente que a **tendência** não dita o **comportamento** da distribuição, depende inclusive do tipo de dado que está sendo analisado.

• Explicações dos Casos de Distribuição -

(2.1)

Caso queira pular a explicação, o próximo tópico começa a partir da página 13.

Gráfico 2.11



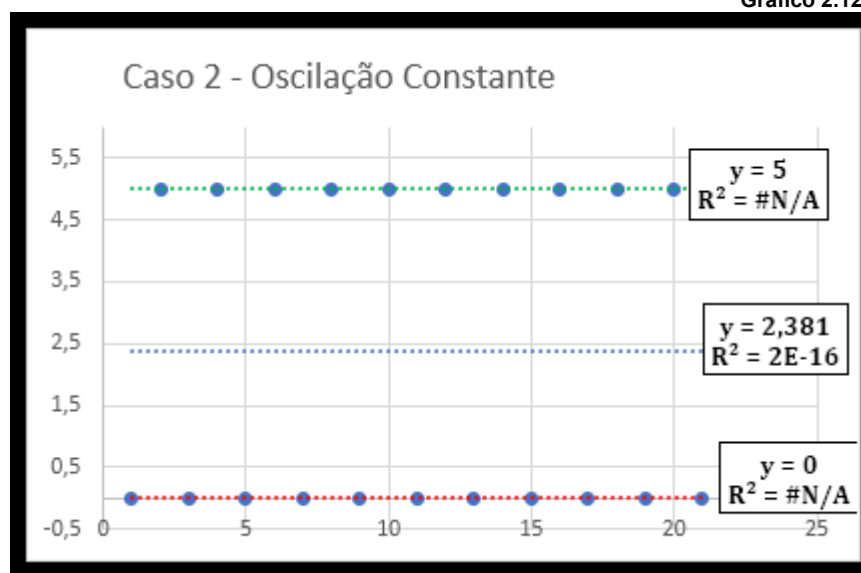
- **Caso 1 - Ajuste Linear Padrão**

Os casos “1” são a forma mais comum e utilizada de ajuste linear, buscando uma relação de **comportamento** já presente e determinado com base em outros cálculos, como por exemplo, um gráfico de “posição x tempo” determinando um “movimento retilíneo uniforme”. Observe que a *correlação* é alta, pois ela varia conforme a distância dela pros pontos. Essa alta *correlação* costuma indicar nesses casos uma variação na forma de medição, erro humano ou questões de natureza parecida. Observe também que o caso 1.3 é um ajuste onde o ponto inicial está definido previamente em relação ao cálculo de regressão linear.

- **Caso 2 - Oscilação Constante**

Nos casos “2” se pode observar que as correlações são baixíssimas (inclusive nesses casos especificamente quanto menor mais preciso). Mesmo tendo uma baixa *correlação* a **tendência** dita pelo ajuste linear global se ajusta bem ao **comportamento** da oscilação dos dados, pois eles oscilam de forma constante, tendo tanto seus “*dados inferiores*” quanto seus “*dados superiores*” constantes.

Gráfico 2.12

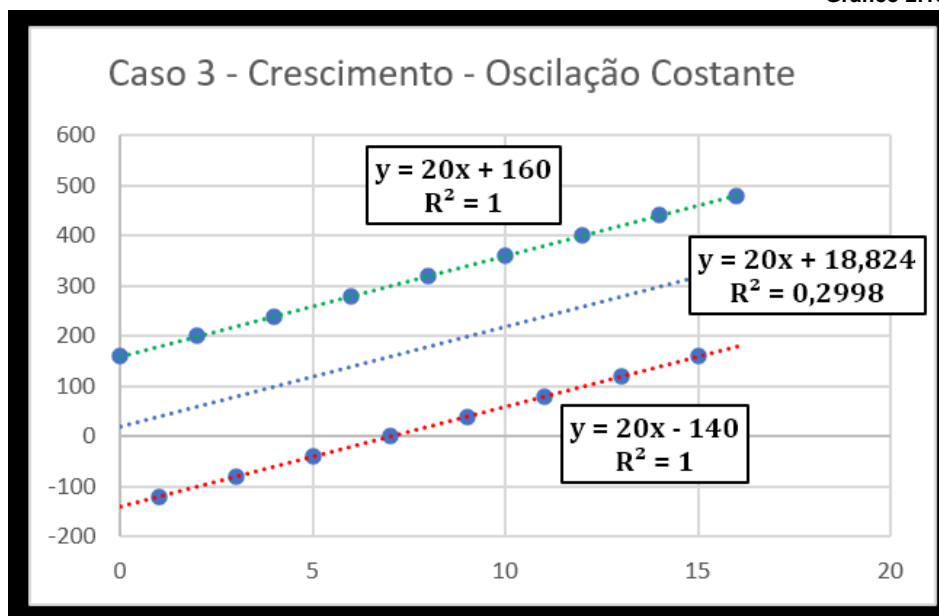


Pode-se concluir que apesar do encaixe e correlação não serem muito precisos, a reta dita muito bem o **comportamento** da oscilação (ser absolutamente constante, ou seja, tem apenas o coeficiente linear, não tendo um coeficiente angular).

- **Caso 3 - Oscilação Constante em Crescimento**

Nos casos “3” se pode observar que as correlações são baixíssimas. Mesmo tendo uma baixa *correlação* a **tendência** dita pelo ajuste linear se ajusta bem ao **comportamento** da oscilação dos dados, pois eles oscilam de forma constante mesmo em crescimento, tendo tanto seus “*dados inferiores*” quanto seus “*dados superiores*” em crescimento constante.

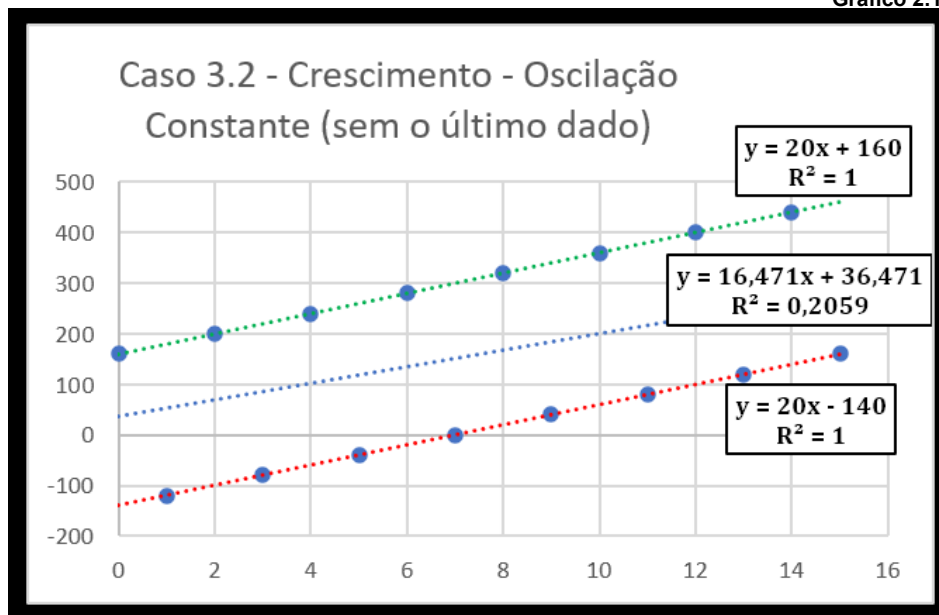
Gráfico 2.13



A partir da análise deste tipo de caso, podemos observar que apesar da correlação ser de $\approx 30\%$, o ajuste linear global dita perfeitamente o **comportamento** de crescimento desta oscilação, tendo o exato mesmo coeficiente angular (**20x**) para os “**dados superiores**”, “**dados inferiores**” e também para o ajuste linear total.

- **Casos 2 e 3 - Falta ou Anomalia de Dados**

Gráfico 2.14



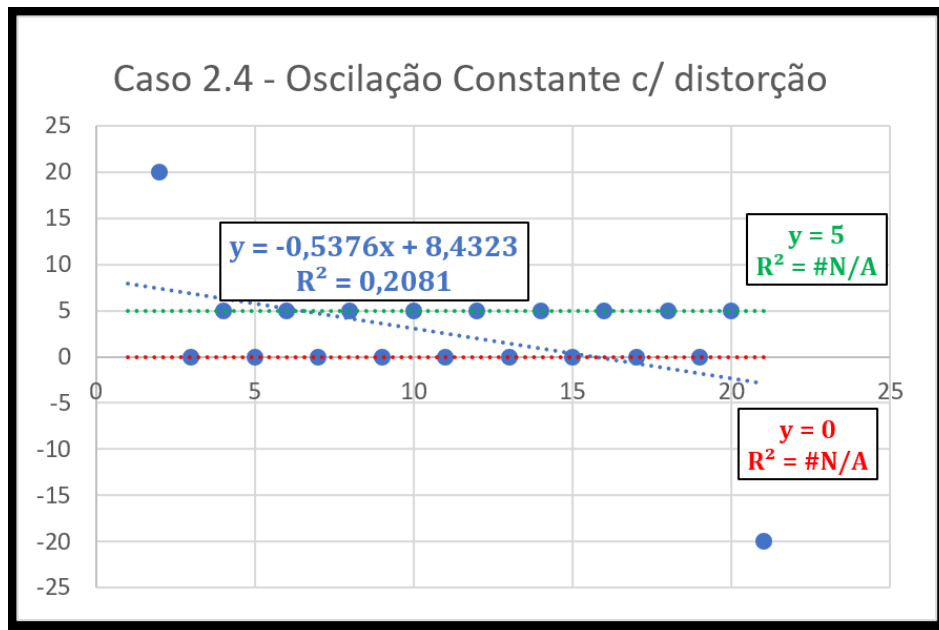
Observe como a falta de 1 único dado distorce o ajuste, mesmo os “**dados superiores**” e “**dados inferiores**” estando perfeitos, o ajuste linear global não está, independente da correlação ser boa ou não, o coeficiente angular teve uma discrepância relativa de $\approx 17,65\%$ e absoluta de 3,529.

O mesmo acontece nos casos 2.1, 2.2 e 2.3, onde o ajuste da oscilação constante passa a ter um coeficiente angular, deixando de ser uma constante.

O Coeficiente Angular se mostra próximo de 0, mas no caso 3.2 um único dado foi capaz de gerar uma discrepância relativa de $\approx 17,65\%$. Como a ideia do estudo é buscar se o **comportamento** é de crescimento, queda ou constância, essas distorções não alteram o ajuste linear global o suficiente para mudar o **comportamento**, apesar de ser gerado um dado com uma precisão menor mas ainda com uma boa acurácia. Apenas uma análise qualitativa visual já será o suficiente na maior parte dos casos.

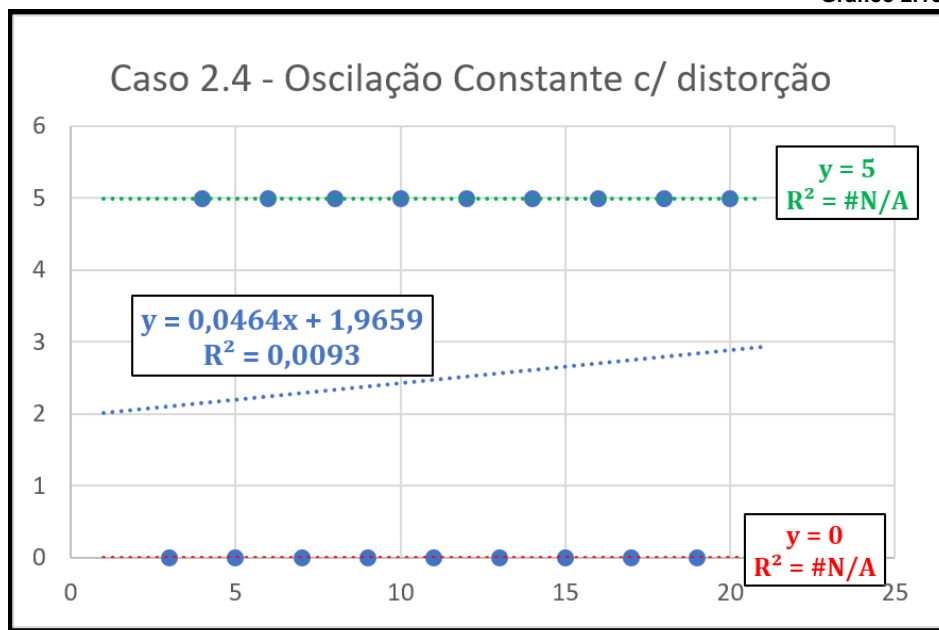
O Caso 2.4 se mostra o único capaz de alterar completamente a leitura de um **comportamento**, com dados “anormais” nas extremidades a leitura de uma constância com oscilação nas extremidades (olhando para o ajuste) passaria a ser de um decréscimo.

Gráfico 2.15



Excluindo os dados das pontas retornamos ao caso 2.1, só que potencializado pois agora não temos nem o último nem o primeiro dado.

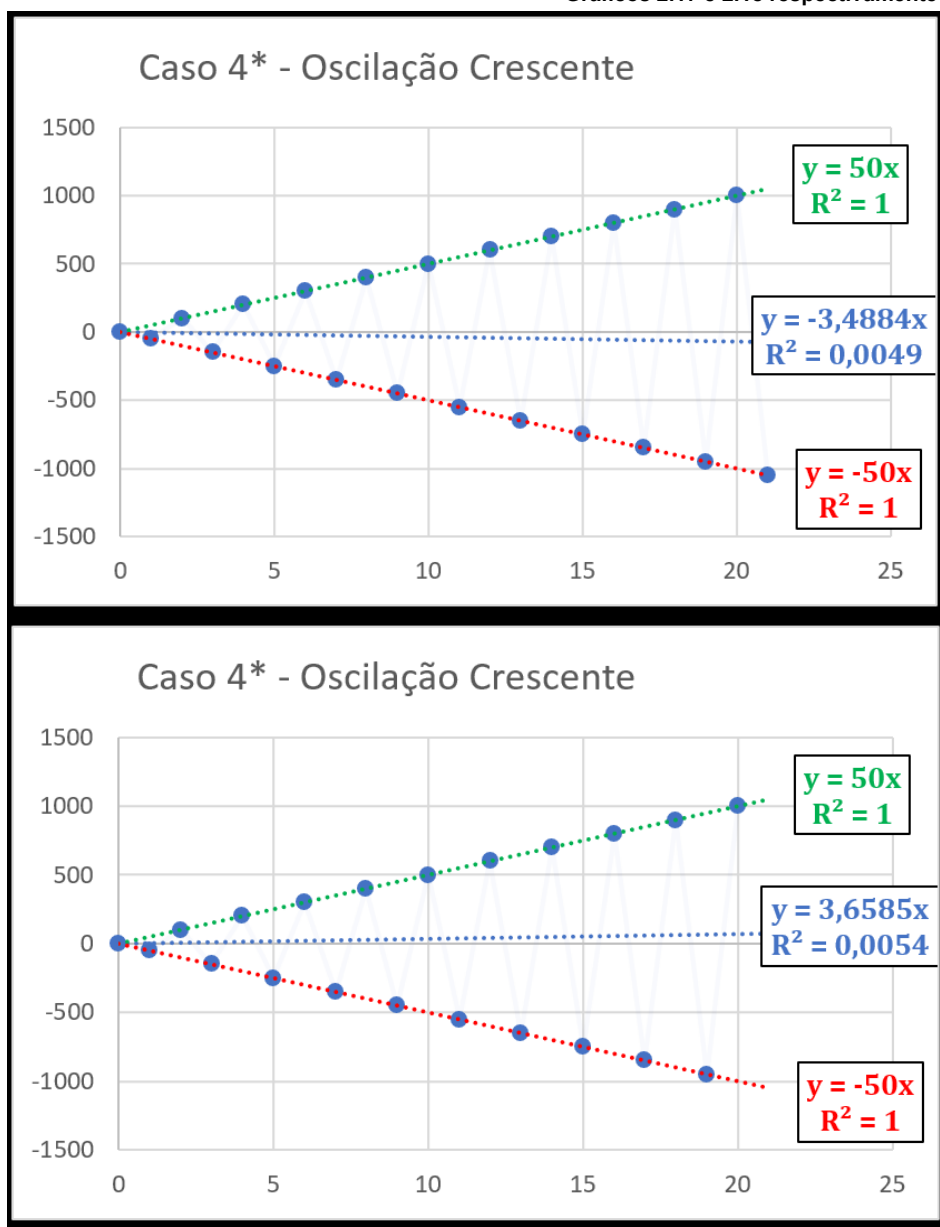
Gráfico 2.16



Não é nada que uma análise visual não possa detectar (uma oscilação constante).

● Caso 4 - Oscilação Crescente

Gráficos 2.17 e 2.18 respectivamente



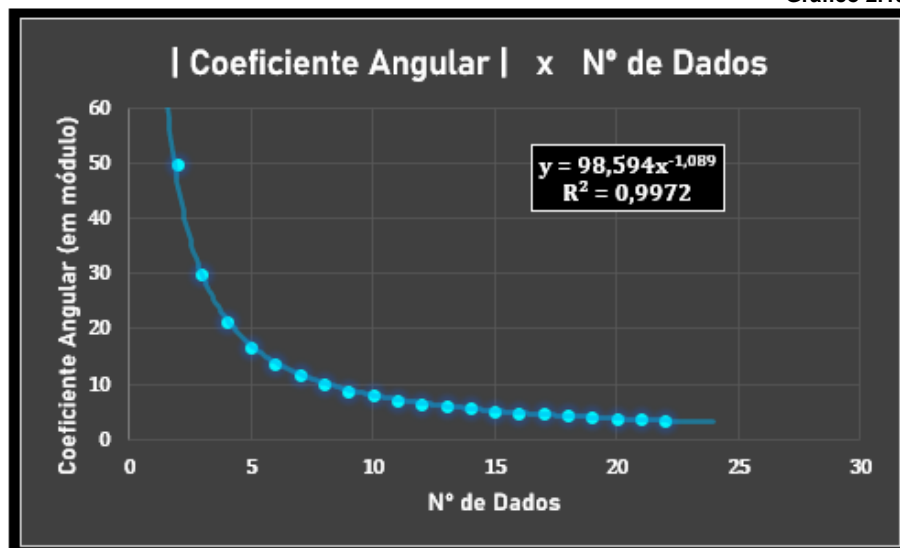
É possível observar que a **tendência** da reta varia conforme qual é o último dado (diferença dos 2 gráficos ao lado).

Também é possível averiguar que quanto mais dados oferecidos, mais próxima de 0 foi a reta do ajuste global.

Gráfico 2.19

Nº de Dados → Coeficiente Angular

6 dados → $\approx -13,636x$
 8 dados → $\approx -10,000x$
 12 dados → $\approx -6,5217x$
 15 dados → $\approx -5,1724x$
 18 dados → $\approx -4,2857x$
 21 dados → $\approx -3,6585x$



A oscilação cresce de forma constante (e igual nesse caso) tanto nos “**dados superiores**” quanto nos “**dados inferiores**”, assim, a soma de seus coeficientes angulares é 0, que seria a melhor reta de ajuste linear possível ($y=0$) observando que a reta interior vai se aproximando de 0 quanto mais dados são oferecidos ao ajuste.

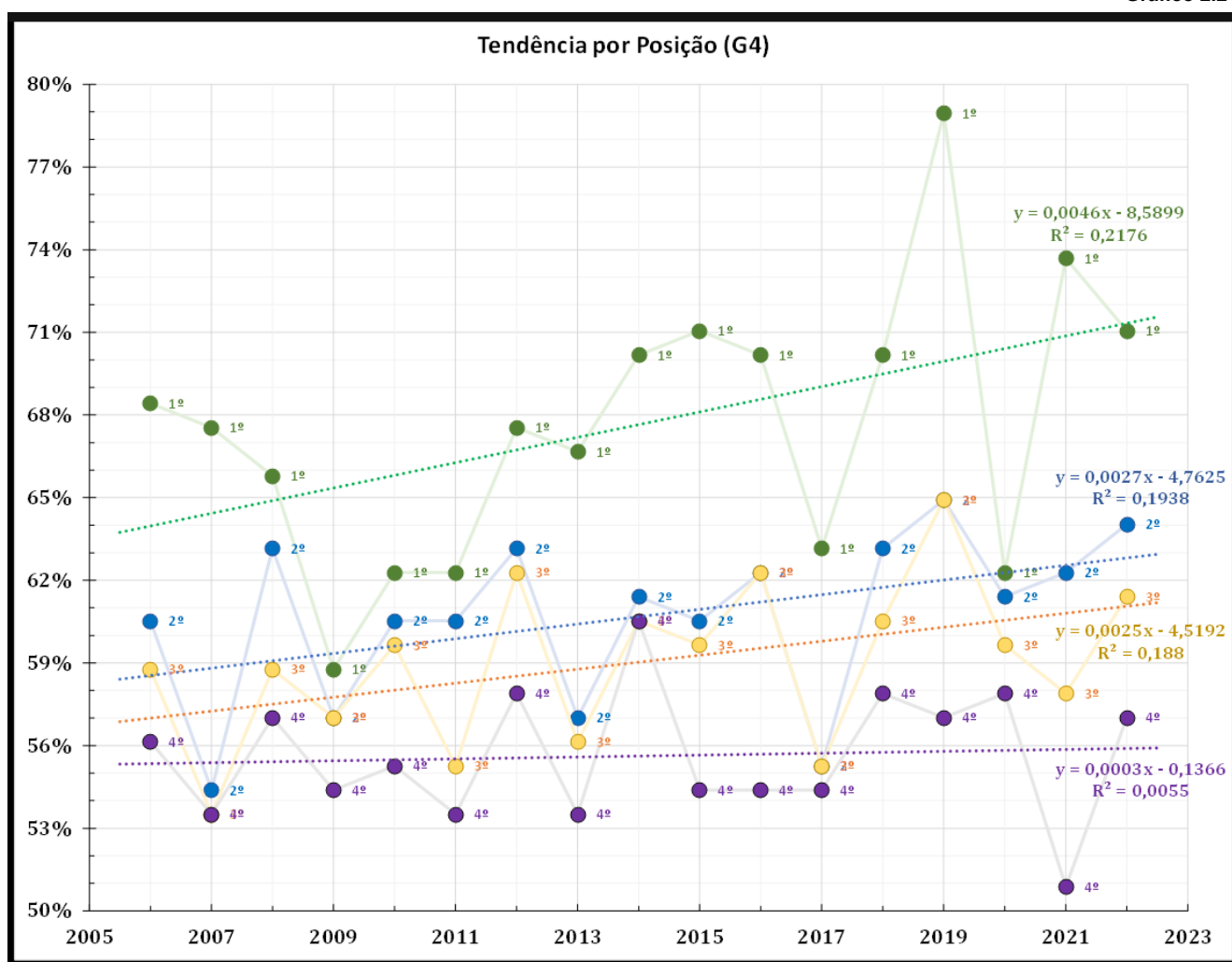
• Continuação da Análise

(2.2)

Todas essas explicações acima servem para embasar a análise qualitativa que será realizada nessa seção (análise 2).

Para esta sequência, como será observado o comportamento das 20 posições da tabela, não há como inserir os dados das pontuações dos anos 2003, 2004 e 2005, então seguirá a partir de 2006 até 2022.

Gráfico 2.21



Como se pode observar, do 1º ao 4º, há um crescimento dos “**dados superiores**” e “**dados inferiores**”. O coeficiente angular dessas retas significa supostamente o quanto o aproveitamento tem crescido por ano. 0,0046 significaria que cresce $\approx 0,46\%$ ao ano o que em 10 anos seria $\approx 4,6\%$ que em pontos se traduz em $\approx 5,25$ pontos. O quarto colocado tem uma distorção causada pelas oscilações de 2014 e 2021, pois tanto seus “**dados superiores**” e “**dados inferiores**” tem crescido, o que indica uma tendência de crescimento, apesar do ajuste linear global do 4º colocado indicar estabilidade.

Para se certificar que a leitura dos dados está sendo feita corretamente, faz-se necessário a análise qualitativa citada anteriormente, desconsiderando por exemplo, as maiores oscilações principalmente do começo e fim.

Considerar **Platô** e **Planície** a **faixa de dados superiores e inferiores** (respectivamente).

A análise completa é feita a partir de algumas etapas:

- **Etapas 1:**

- Plotar os gráficos dos aproveitamentos por ano para todas as posições e realizar o Ajuste/Regressão Linear. (1.1)
- Observar visualmente o **comportamento**, observar se os dados “**inferiores**” e “**superiores**” fazem sentido com o ajuste linear. (1.2)
- Descrever a **tendência do comportamento** analisando apenas visualmente e o coeficiente angular do ajuste. (1.3)

- **Etapas 2:** (estarão definidos por LFC+)

- Utilizar os gráficos da etapa 1 para as próximas partes. (2.1)
- Observar visualmente as maiores oscilações distorções que ultrapassam o **platô** e o **planície** principalmente nas **pontas**. (2.2)
- Apagar as maiores oscilações das pontas, e observar a **tendência do comportamento** e ajuste linear. Os critérios para a escolha das oscilações estão definidos abaixo (gerando novos gráficos). (2.3)

- **Etapas 3:** (estarão definidos por LFC++)

- Utilizar os gráficos da etapa 2 para as próximas partes. (3.1)
- Observar visualmente as maiores oscilações distorções que ultrapassam o **platô** e o **planície** ao longo do gráfico independente da localização. (3.2)
- Remover as maiores oscilações, os critérios estão definidos abaixo (gerando novos gráficos). (3.3)
- Descrever a **tendência do comportamento** analisando **apenas** os LFC+ e LFC++ e seus respectivos “Fit”. (3.4)

- **Etapas Final:**

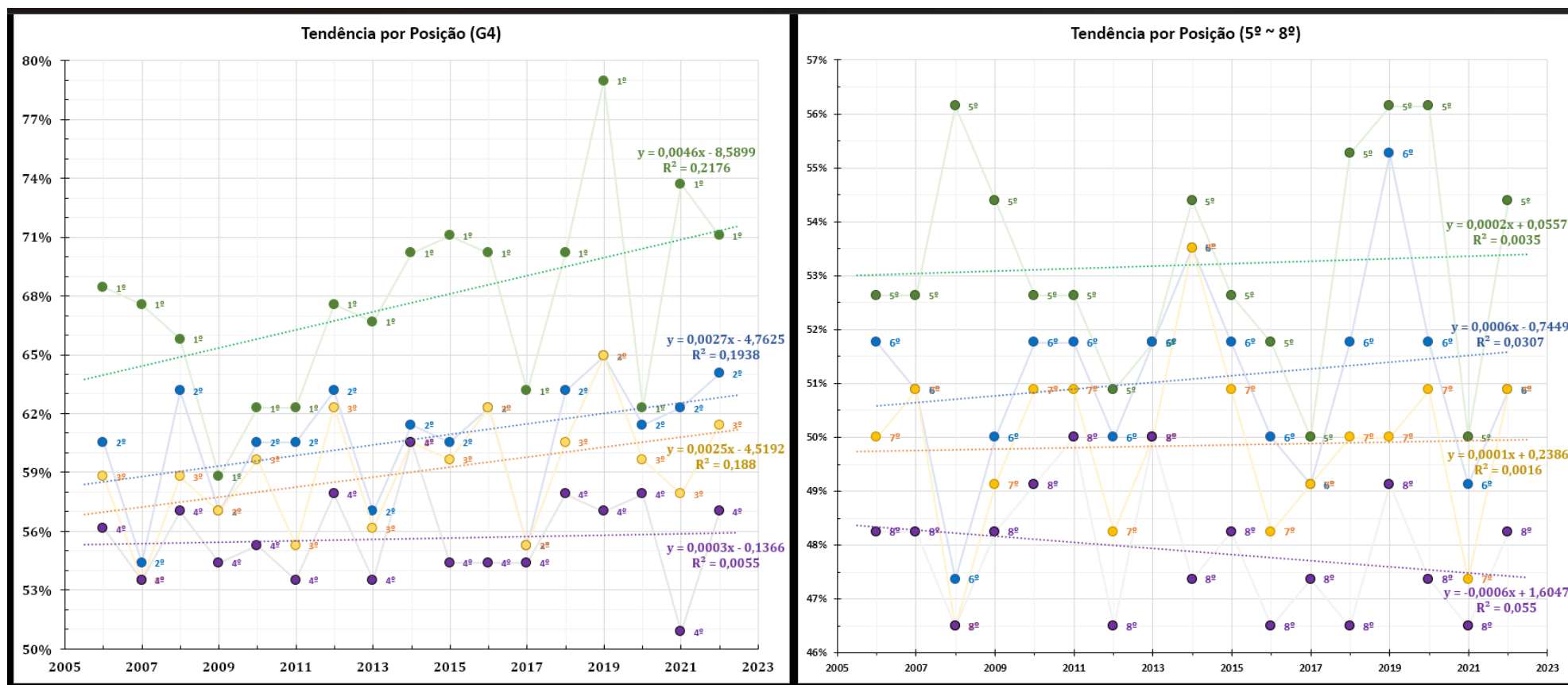
- Descrever a **tendência do comportamento** analisando o conjunto total de informações adquiridas e separar quais seriam o blocos de cresc. e decresc.

- **Critérios:**

- Para realizar essa remoção do dado, precisa se aproximar suficientemente ou superar 2,5% de oscilação no aproveitamento, não apenas de um ano pro outro como também se ultrapassa o **platô** e a **planície** local.
- Oscilações menores podem ser apagadas em caso de oscilações constantes com menos dados ou em casos de ter perfurado o **platô/planície** (casos explicados anteriormente)
- Ajustar ao máximo possível entre oscilação positiva e negativa (para não apagar todas negativas ou todas positivas e gerar uma nova distorção).

- **Etapa 1**

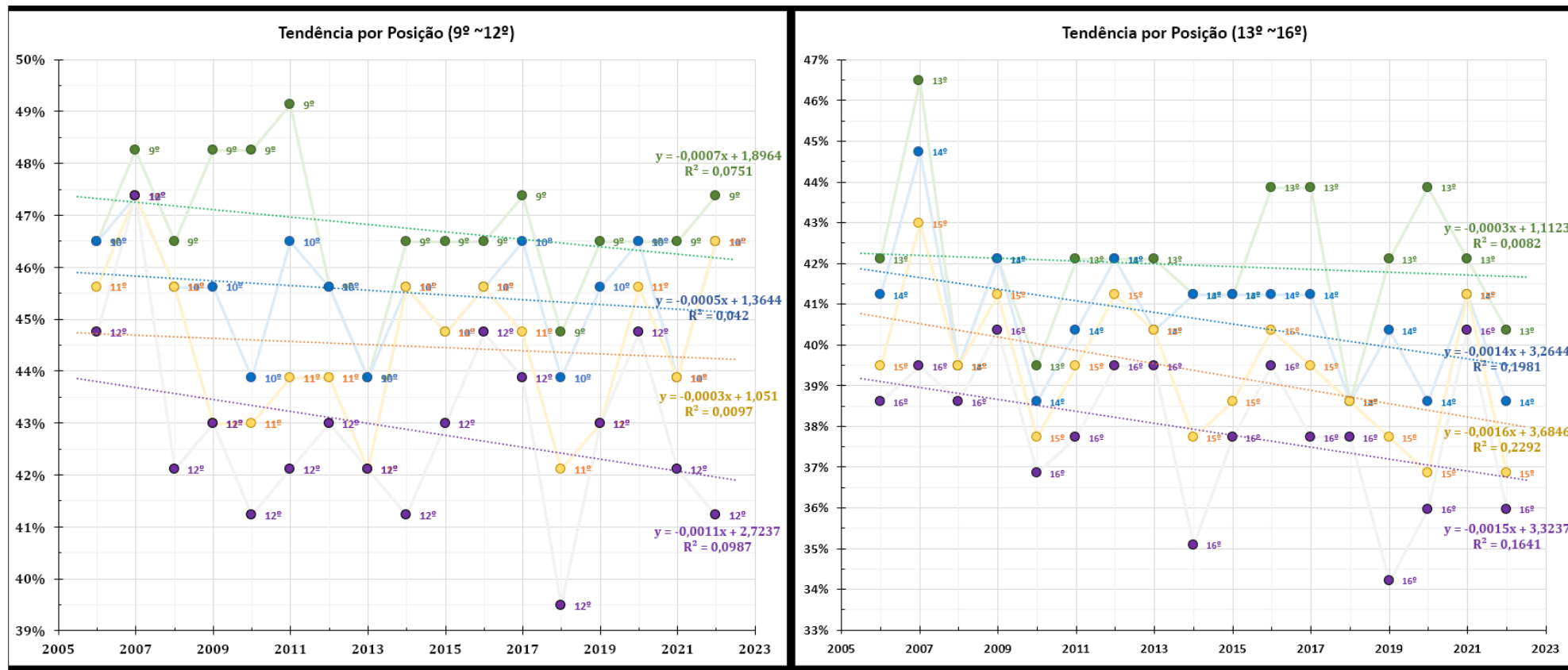
Gráficos do G4 (1º ao 4º) e do 5º ao 8º



Gráficos 2.2.11 e 2.2.12

- **Etapa 1**

Gráficos do 9º ao 12º e do 13º ao 16º

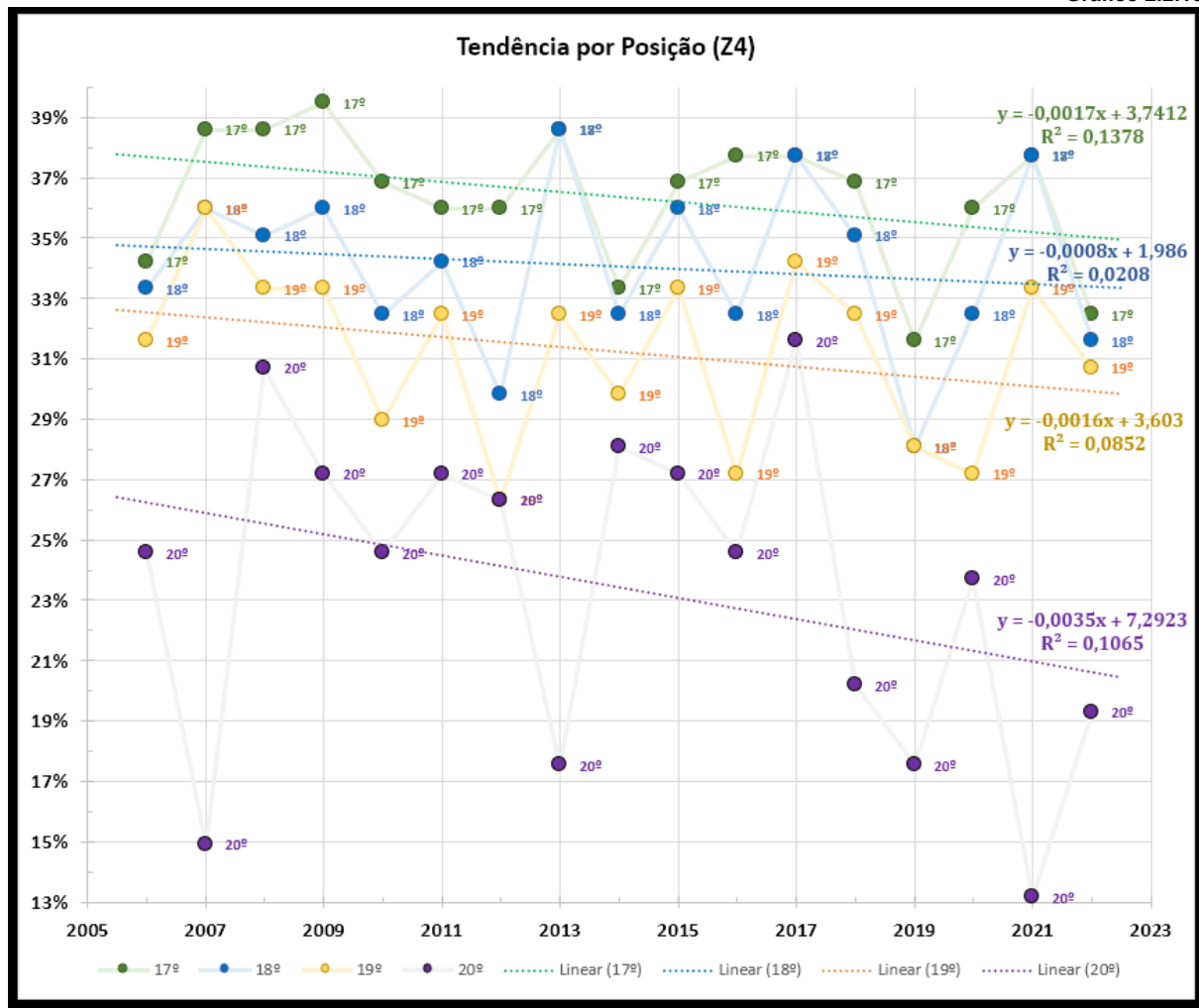


Gráficos 2.2.13 e 2.2.14

• Etapa 1

Gráfico do Z4 (17º ao 20º)

Gráfico 2.2.15

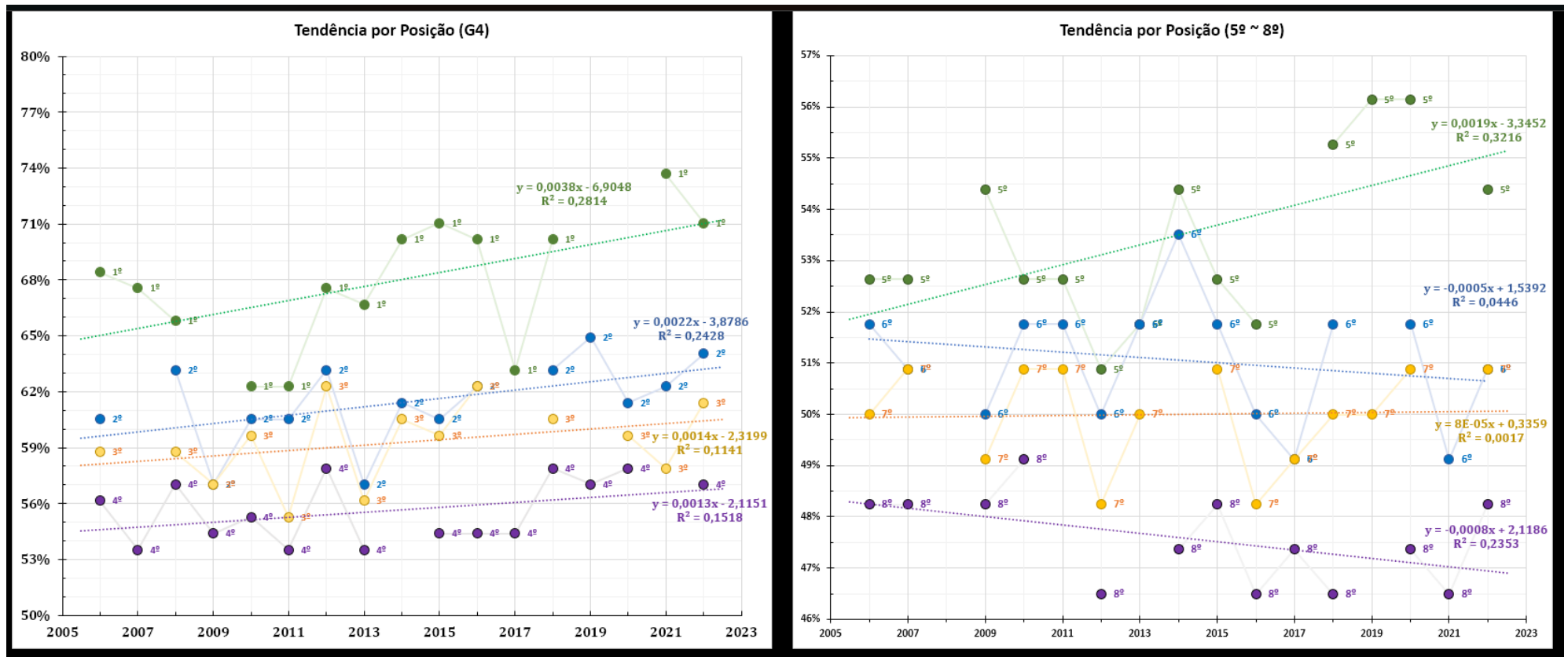


1.2 e 1.3

Pos	Méd pts	Méd Apr	LFC	Encaixe	Análise Qualitativa Primária do Gráfico	LFC - Coeficiente Regressão Linear	Tabela 2.2.16
1	77,1	67,6%	0,46%	21,8%	1 Crescimento muito acelerado		
2	69,2	60,7%	0,27%	19,4%	2 Crescimento acelerado		
3	67,3	59,0%	0,25%	18,8%	3 Crescimento acelerado		5º - a falsa estabilidade se dá por uma positiva no começo e negativa no fim.
4	63,4	55,6%	0,03%	0,55%	4 Cresce Pouco		
5	60,6	53,2%	0,02%	0,35%	5 Crescimento das Oscilações*		6º - O Crescimento se dá pela oscilação negativa no começo e positiva no fim.
6	58,2	51,1%	0,06%	3,07%	6 Estável + Crescimento das Oscilações positivas		Mas ainda parece estar Diminuindo as oscilações negativas e aumentando as oscilações positivas
7	56,8	49,8%	0,01%	0,16%	7 Planície Oscilando, Platô Estável (1 pico)		11º - 2 picos e 2 vales fora da oscilação
8	54,6	47,9%	-0,06%	5,50%	8 Platô Decresce com a Planície Estável		
9	53,3	46,7%	-0,07%	7,51%	9 Platô Decr. Planície Oscila, Meio Estável		
10	51,9	45,5%	-0,05%	4,20%	10 Platô e Planície Estáveis. Apenas 1 Osc. Fora		12º - O decréscimo se dá por 2 oscilações, uma positiva no começo e uma negativa ao fim. Há 2 grupos de dados principais q indicaria o crescimento.
11	50,7	44,5%	-0,03%	0,97%	11 Oscila muito (decresceu e agora cresce)*		
12	48,9	42,9%	-0,11%	9,87%	12 Não corresponde ao gráfico*		
13	47,8	42,0%	-0,03%	0,82%	13 Distorção nas pontas + Aum. da Oscilação		Cresceu, decresceu e cresceu novamente
14	46,4	40,7%	-0,14%	19,8%	14 Platô e Planície Decrescem		
15	44,9	39,4%	-0,16%	22,9%	15 Platô e Planície Decrescem		13º - 2 dados nas pontas que causam distorção, picos de oscilação e cresc. do Platô e Decresc da planície
16	43,2	37,9%	-0,15%	16,4%	16 Platô e Planície Decrescem		
17	41,5	36,4%	-0,17%	13,8%	17 Platô e Planície Decrescem		
18	38,8	34,1%	-0,08%	2,08%	18 Platô e Planície Decrescem Pouco		
19	35,6	31,2%	-0,16%	8,52%	19 Platô e Planície Decrescem (com picos cada vez menores)		
20	26,7	23,4%	-0,35%	10,7%	20 Platô e Planície Decrescem Muito (Com Osc.)		

● Etapa 2

Gráficos do G4 (1º ao 4º) e do 5º ao 8º



Gráficos 2.2.21 e 2.2.22

- Etapa 2

Gráficos do 9º ao 12º e do 13º ao 16º

Gráficos 2.2.23 e 2.2.24



● Etapa 2

Gráfico do Z4 (17º ao 20º)

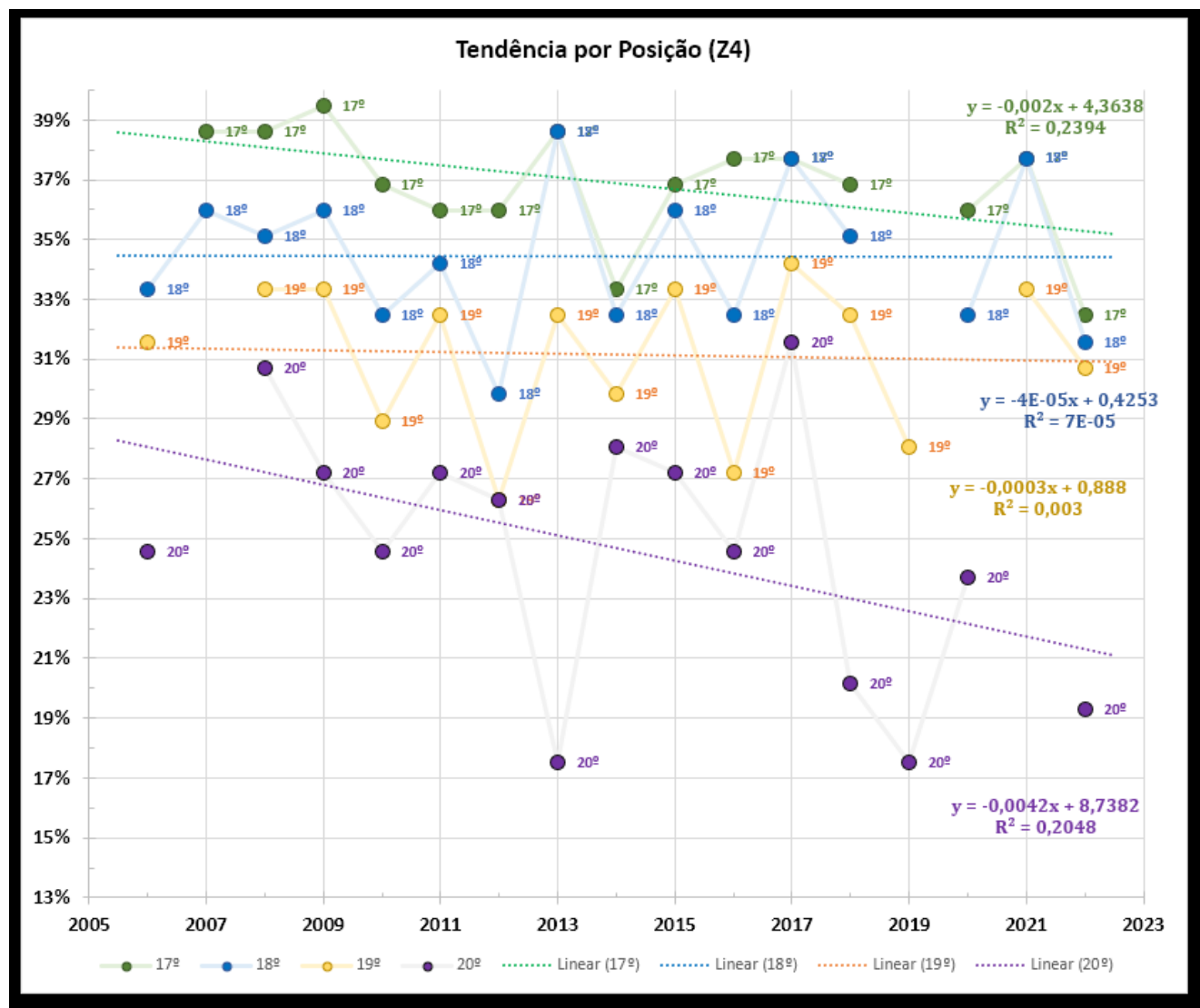
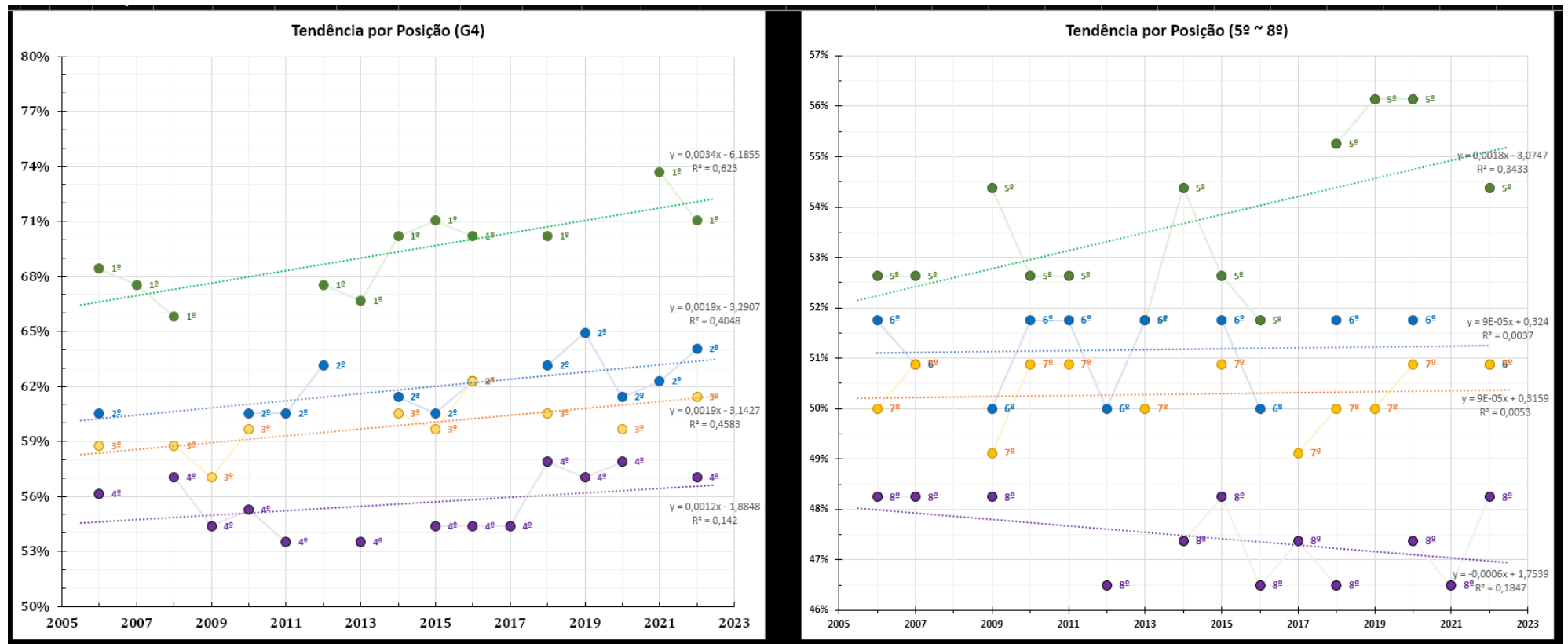


Gráfico 2.2.25

● Etapa 3

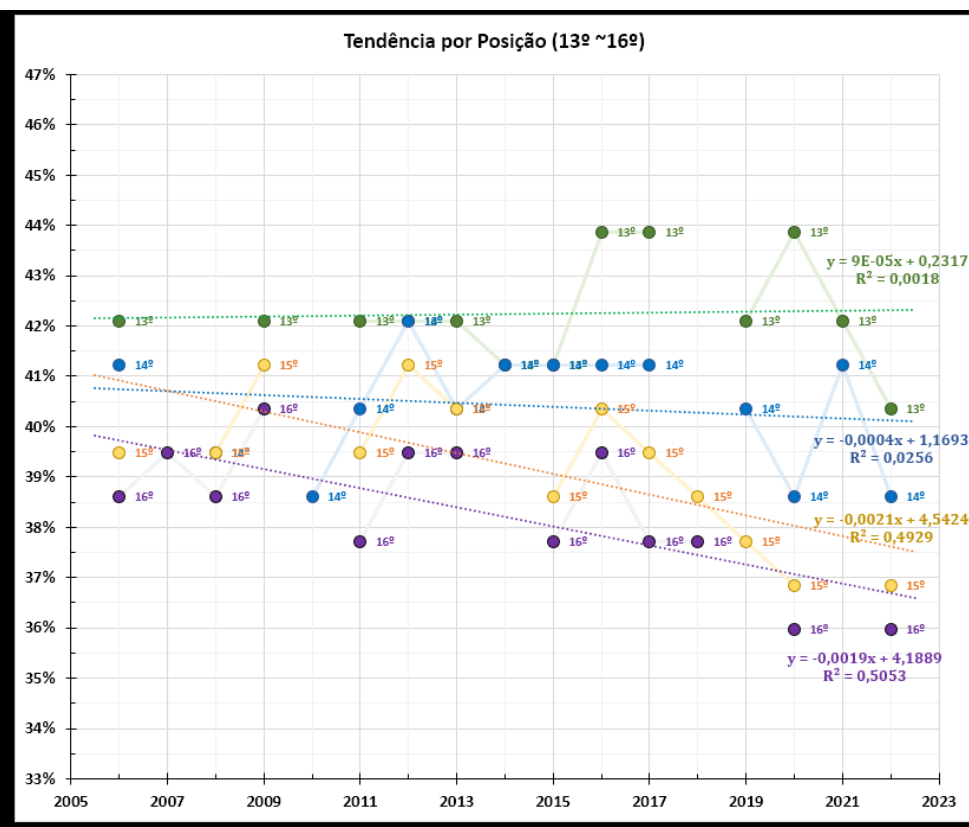
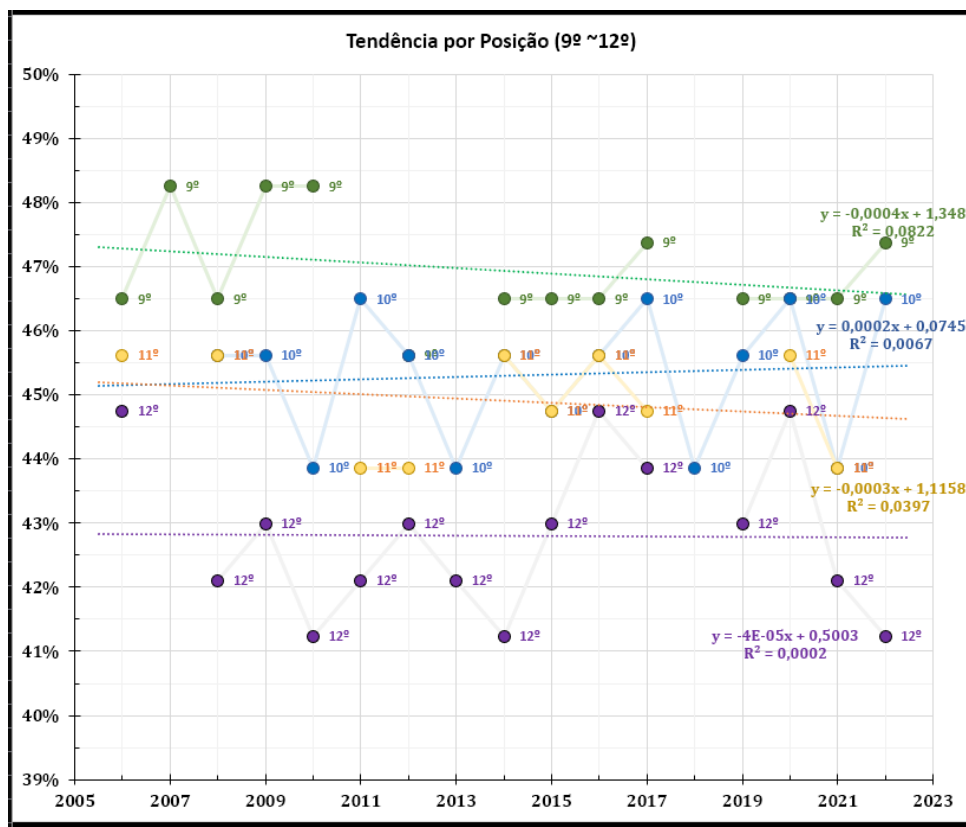
Gráficos do G4 (1º ao 4º) e do 5º ao 8º



Gráficos 2.2.31 e 2.2.32

- Etapa 3

Gráficos do 9º ao 12º e do 13º ao 16º



Gráficos 2.2.33 e 2.2.34

• Etapa 3

Gráfico do Z4 (17º ao 20º)

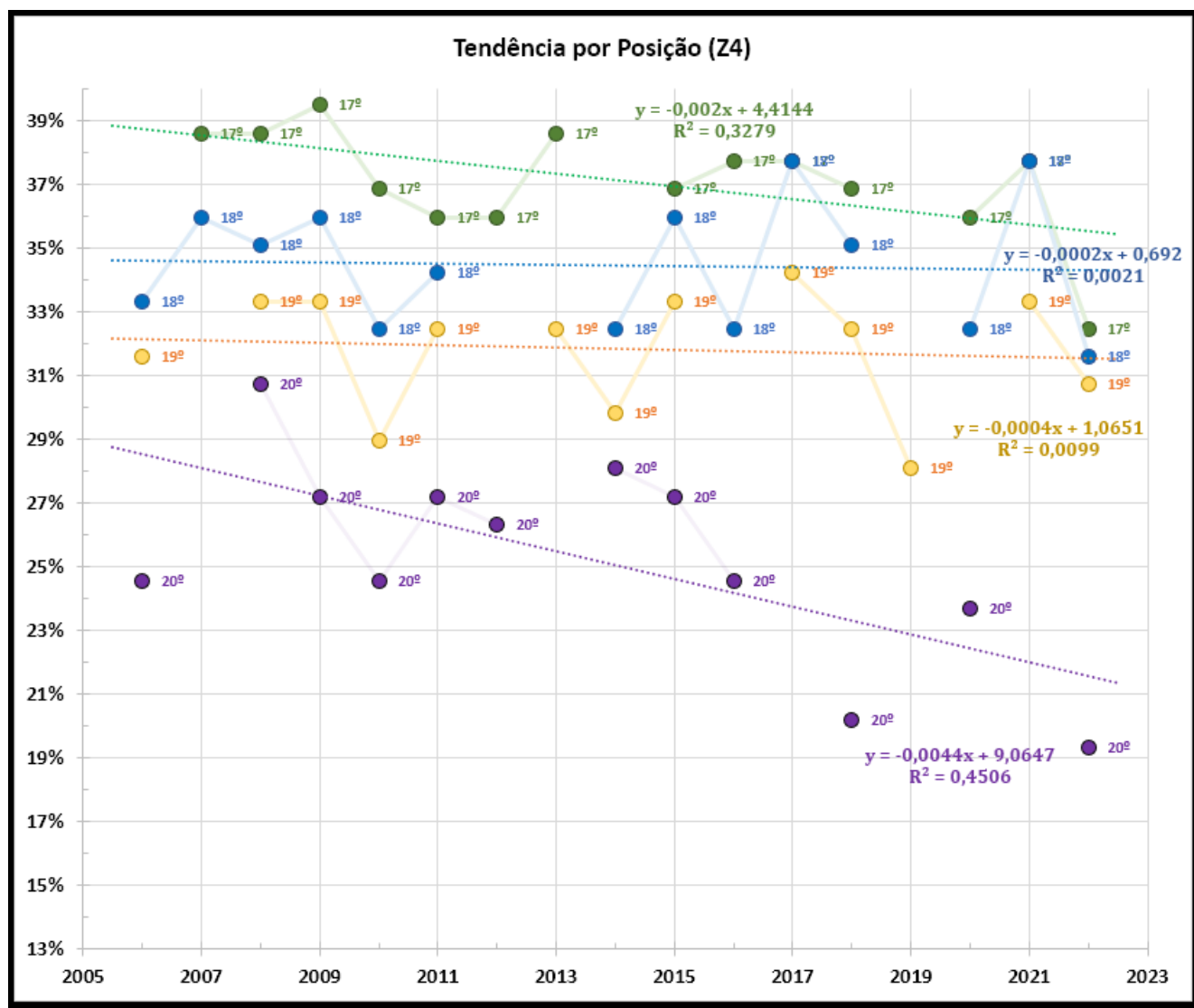


Gráfico 2.2.25

Remoções:

Pos	Etapa 2	Etapa 3
1º	09, 19, 20	09, 10, 11, 17, 19, 20
2º	07, 17	07, 08, 09, 13, 17
3º	07, 17, 19	07, 11, 12, 13, 17, 19, 21
4º	14, 21	07, 12, 14, 21
5º	08, 17, 21	08, 12, 17, 21
6º	08, 19	08, 14, 17, 19, 21
7º	08, 14, 21	08, 12, 14, 16, 21
8º	08, 11, 13, 19	08, 10, 11, 13, 19
9º	11, 13, 18	11, 13, 18
10º	07	06, 07

Pos	Etapa 2	Etapa 3
11º	07, 18	07, 09, 10, 13, 18, 19, 22
12º	07, 18	07, 18
13º	07, 10, 18	07, 08, 10, 18
14º	07	07, 09, 18
15º	07, 21	07, 10, 14, 21
16º	19, 21	10, 14, 19, 21
17º	06, 19	06, 14, 19
18º	19	12, 13, 19
19º	07, 20	07, 12, 16, 20
20º	07, 21	07, 13, 17, 19, 21

tabela 2.2.26

• Etapa 3

Pos	LFC+	Fit+	LFC++	Fit++	Análise APENAS do LFC+ e LFC++
tabela 2.2.27					
1	0,38%	28,1%	0,34%	62,3%	1 Crescimento muito acelerado
2	0,22%	24,3%	0,19%	40,3%	2 Crescimento acelerado
3	0,14%	11,4% *	0,19%	45,8%	3 Crescimento acelerado
4	0,13%	15,2%	0,12%	14,2%	4 Cresce consistentemente
5	0,19%	32,2%	0,18%	34,3%	5 Crescimento acelerado *
6	-0,05%	4,46%	0,01%	0,37%	6 Osc. cres. + Tend. de estabilidade
7	0,01%	0,17%	0,01%	0,53%	7 Oscilação em tend. de cresc.
8	-0,08%	23,5%	-0,06%	18,5%	8 Decresce
9	-0,04%	8,22%	-0,04%	8,22%	9 Decresce
10	-0,01%	0,37% *	0,02%	0,67%	10 "Estabilidade"
11	0,05%	3,73%	-0,03%	3,97%	11 Cresce Pouco* + Aum da Osc
12	-0,004%	0,02%	-0,004%	0,02%	12 "Estabilidade"
13	0,06%	6,34%	0,01%	0,18%	13 Estável com picos de oscilação
14	-0,08%	8,98%	-0,04%	2,56%	14 Declínio
15	-0,18%	32,9%	-0,21%	49,29%	15 Declínio acelerado
16	-0,18%	30,1%	-0,19%	50,53%	16 Declínio acelerado
17	-0,20%	23,9%	0,20%	32,79%	17 Declínio acelerado
18	-0,004%	0,01% *	-0,02%	0,21%	18 Aum. da Osc. + Tend. De Queda
19	-0,03%	0,30% *	-0,04%	0,99%	19 Osc. Estável + Decrésc. do Platô e Vale
20	-0,42%	20,5%	-0,44%	45,06%	20 Declínio muito acelerado

• Etapa Final

Pos	Análise Qualitativa FINAL do Gráfico	Análise dos Blocos
tabela 2.2.28		
1º	Crescimento muito acelerado	Bloco em Crescimento (Acelerado ou não)
2º	Crescimento acelerado	
3º	Crescimento acelerado	
4º	Cresce consistentemente	
5º	Tend. de cresc. é leve pelo aum. das oscilações neg.	Bloco Estável (ou Mirando Estabilidade)
6º	Dim. das osc. neg. + Aum. das osc. pos.	
7º	ESTABILIDADE + Aum. das osc.	
8º	Decresce mirando estabilidade*	
9º	Decresce mirando estabilidade*	Bloco em Declínio (Acelerado ou não)
10º	ESTABILIDADE do Apr e das Osc	
11º	ESTABILIDADE + Aum. das osc.	
12º	Estabilidade nas faixas e Osc. Estável	
13º	Estável com picos de oscilação	
14º	Declínio	
15º	Declínio acelerado	
16º	Declínio acelerado	
17º	Declínio acelerado	
18º	Decresce	
19º	Declínio	
20º	Declínio muito acelerado	

• Conclusão

(3.0)

Como se pode ver, há uma tendência clara de crescimento de aproveitamento das posições do G6, estabilidade na faixa intermediária (7º ao 13º) e decréscimo do aproveitamento nas posições do Z7 (independente dos clubes).

Realizando um gráfico com o aproveitamento médio dos blocos (por exemplo, g6, $\frac{\text{(pontuação acumulada G6)}}{114 * 6}$) podemos observar claramente essa tendência:

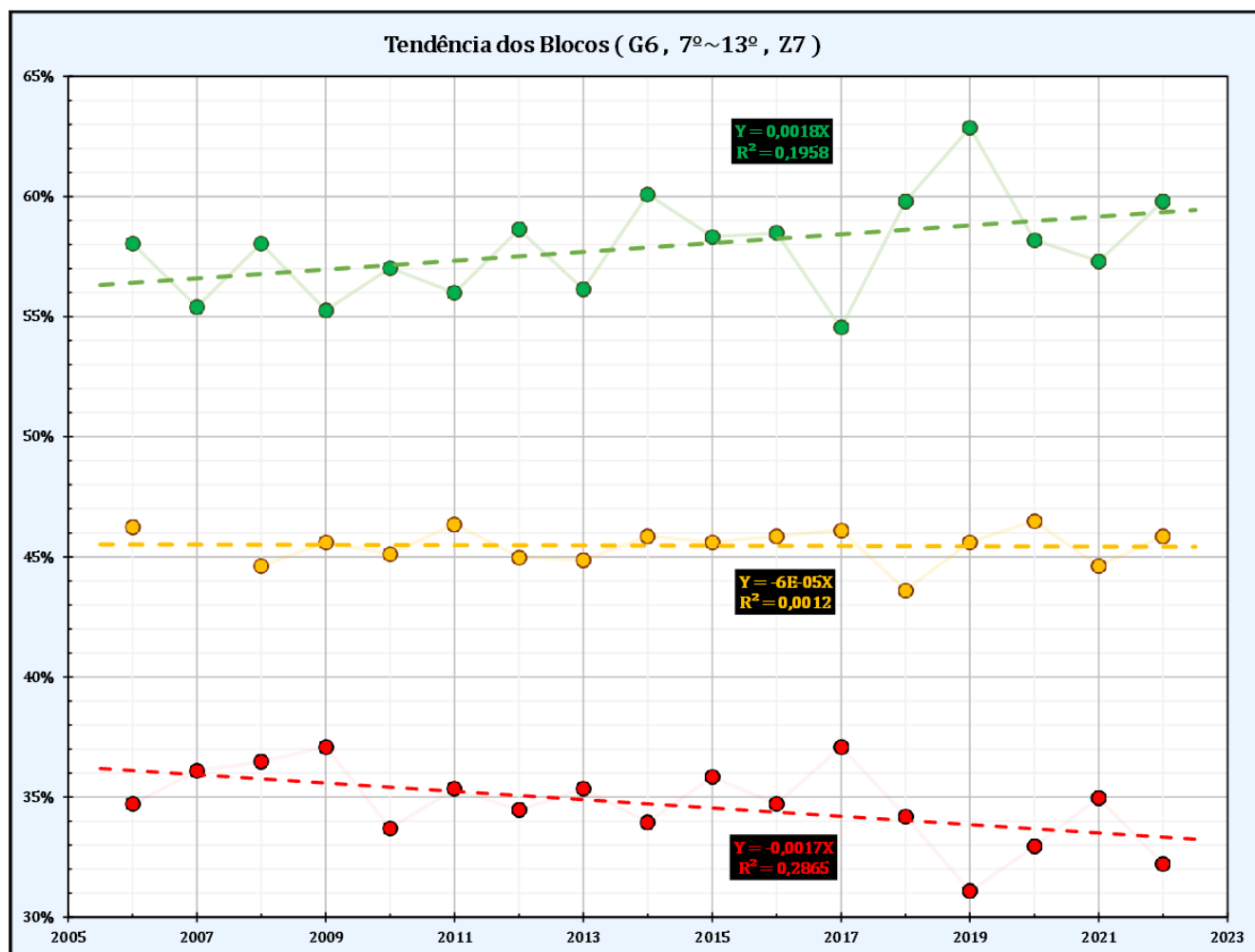


gráfico 3.01

A soma dos **coeficientes angulares** das retas chegar próximo de 0 faz todo sentido, uma vez que a pontuação total do campeonato só pode variar entre **760** e **1140** dependendo do nº de empates (são 380 jogos, empate é 1 ponto para cada, logo 2 pontos, e vitória é 3 e 0, ou seja, 3 pontos).

Ou seja, é uma oscilação com **Platô** e **Planície** já bem definidos e o que varia nesse caso é a **DISTRIBUIÇÃO** desses pontos nessas oscilações. (**caso 2 das explicações**).

Coeficiente Angular da reta é uma relação $\frac{\Delta Y}{\Delta X}$, que diz o quanto subiu no eixo y para cada unidade do eixo x. A ideia que se tem analisando os blocos é que o G6 cresceu cerca de 3% em 17 anos (extremidades do segmento da reta), o que seria nesse caso cerca de 0,1765% ao ano. Um exemplo aproximado para entender o quanto vale esse crescimento. o G6 aumentar 0,18% ao ano quer dizer que em média a cada 10 anos, os clubes em média fizeram 2 pontos a mais. Ou seja, 20 anos, 4 pontos e assim sucessivamente.

Conversão de APR para Pts				
tabela 3.02				
Pos	Méd pts	LFC	LFC+	LFC++
1º	77,1	0,52	0,43	0,39
2º	69,2	0,31	0,25	0,22
3º	67,3	0,29	0,16	0,22
4º	63,4	0,03	0,15	0,14
5º	60,6	0,02	0,22	0,21
6º	58,2	0,07	-0,06	0,01
7º	56,8	0,01	0,01	0,01
8º	54,6	-0,07	-0,09	-0,07
9º	53,3	-0,08	-0,05	-0,05
10º	51,9	-0,06	-0,01	0,02
11º	50,7	-0,03	0,06	-0,03
12º	48,9	-0,13	0,00	0,00
13º	47,8	-0,03	0,07	0,01
14º	46,4	-0,16	-0,09	-0,05
15º	44,9	-0,18	-0,21	-0,24
16º	43,2	-0,17	-0,21	-0,22
17º	41,5	-0,19	-0,23	0,23
18º	38,8	-0,09	0,00	-0,02
19º	35,6	-0,18	-0,03	-0,05
20º	26,7	-0,40	-0,48	-0,50

Mesmo assim, é possível ter uma ideia do quanto cresceu ou decresceu observando a **tabela 3.02** ao lado. Deve-se observar o comportamento do LFC, LFC+ e LFC++. Como por exemplo o 5º e 12º lugares, os quais os comportamentos são bem descritos pelos LFC+ e LFC++. O 6º que oscila muito e deu 3 resultados diferentes, podendo se observar que se cresce pouco.

84 pontos foi a projeção dada pelo cruzamento das retas da **página 7, gráficos 2.02 e 2.03**. O que corrobora toda a análise feita até este ponto.

Já comprovado que as distâncias entre os blocos. Pode-se agora avaliar quais times frequentam mais quais blocos, apesar disso não fazer exatamente parte da comprovação, pode-se gerar outros tipos de conclusões sobre esse tema.

O gráfico abaixo usa como base o histórico completo das classificações dos times da série A:

[illegible]

Nesta imagem acima não dá pra enxergar, mas na fonte está em condições (tanto de onde foi tirado a tabela quanto o arquivo excel utilizado).

tabela 3.04

	Times	Ordem de Pontuação	Verde	Amarelo	Red/ Laranja	Partici pações	Pont uação
1º	São Paulo		13	7	0	20	46
2º	Internacional		12	6	1	19	42
3º	Flamengo		12	4	4	20	40
4º	Corinthians		10	8	1	19	38
5º	Grêmio		11	4	3	18	37
6º	Palmeiras		10	5	3	18	35
7º	Santos		7	12	1	20	33
8º	Atlético-MG		8	8	3	19	32
9º	Cruzeiro		8	7	2	17	31
10º	Athletico-PR		7	9	3	19	30
11º	Fluminense		8	5	7	20	29
12º	Botafogo		3	9	5	17	18
13º	Goiás		3	7	4	14	16
14º	Vasco		3	6	6	15	15
15º	Coritiba		1	5	8	14	8
16º	Paraná		2	2	2	6	8
17º	Figueirense		0	6	5	11	6
18º	Vitória		1	3	6	10	6
19º	Fortaleza		1	3	3	7	6
20º	Sport		1	2	8	11	5
21º	Bahia		0	4	6	10	4
22º	Juventude		1	1	5	7	4
23º	RB Bragantino		1	1	1	3	4
24º	Ponte Preta		0	3	6	9	3
25º	Atlético-GO		0	3	4	7	3
26º	Avai		1	0	6	7	3
27º	Ceará		0	3	4	7	3
28º	São Caetano		1	0	3	4	3
29º	Chapecoense		0	2	5	7	2
30º	América-MG		0	2	3	5	2
31º	Náutico		0	1	4	5	1
32º	Criciúma		0	1	3	4	1
33º	Guarani		0	1	2	3	1
34º	Paysandu		0	1	2	3	1
35º	Grêmio Barueri		0	1	1	2	1
36º	Portuguesa		0	0	3	3	0
37º	Cuiabá		0	0	2	2	0
38º	Santa Cruz		0	0	2	2	0
39º	América-RN		0	0	1	1	0
40º	Brasiliense		0	0	1	1	0
41º	CSA		0	0	1	1	0
42º	Ipatinga		0	0	1	1	0
43º	Joinville		0	0	1	1	0
44º	Santo André		0	0	1	1	0

O critério de pontuação foi feito de um jeito simples: verde 3 pts, amarelo 1 pt e vermelho 0 pt.

Podemos observar que dos clubes grandes, os que estiveram em crises econômicas durante muito tempo estão mais abaixo, como por exemplo Botafogo, Vasco, o Cruzeiro “caindo” apesar de estar em uma boa posição ainda.

Assim como os clubes de maior poder econômico frequentam mais a parte de cima da tabela, como Flamengo, Corinthians, São Paulo, Inter, Grêmio e Palmeiras.

São 44 todos os times que passaram pela série A de 2003 a 2022, alguns com uma boa pontuação, pois as poucas vezes que frequentou conseguiu atingir uma boa posição antes de ter seu declínio, como por exemplo o Paraná Clube.

Outra conclusão que se pode obter, é que dada a realidade onde o clubes de cima da tabela estão fazendo cada vez mais pontos e os debaixo cada vez menos, é o aumento da disparidade

econômica, pois terminar na série A e parte de cima geram mais receita e maiores recursos para planejar o próximo ano e se continuar se mantendo no topo. Além disso, a dificuldade dos clubes vindos da Série B em se manter na A, uma vez que a disparidade econômica dele para pelo menos os 13 primeiros vai ser muito grande, e a disputa para ver quem cai se concentrando do último ao 14º (de 7 times, cairão 4).

● Fontes

(4.0)

- Arquivo Excel Utilizado Pontuações Notáveis Brasileirão A - Série Histórica.xlsx
feito para esse estudo <https://1drv.ms/x/s!AgPh3shhyhBdgRf1KpfYJdjp-0qq>
- UFMG Probabilidades*¹ - Série A – Probabilidades no Futebol
(21/06/2023)*¹ <https://www.mat.ufmg.br/futebol/serie-a/>
- Campeonatos Brasileiros Pontuação e Classificação: [Campeonato Brasileiro](https://www.ogol.com.br/edition.php?id_edicao=162047)
https://www.ogol.com.br/edition.php?id_edicao=162047 (21/06/2023)

*¹ - A UFMG atualiza os números conforme o campeonato vai afunilando, pois os times trocam muitos pontos e o número mágico pode ir caindo, na data de (07/10/2023) o vice tem 45 pontos na 25ª rodada, ou seja, sua pontuação máxima possível é de 84 pontos, para UFMG nesta data, o número mágico é de 82 pontos. A data utilizada nesse estudo era entre a 10ª e 11ª rodada, quando o campeonato está na fase inicial, porém com percurso o suficiente para o número mágico não ser mais de 100 pontos (uma distorção).