

Instituto de Ciências Matemáticas e Computação (ICMC)

Engenharia de Computação

Henrique de Oliveira Araujo - 13863950

Felipe Aparecido da Silva - 11954502



Estatística Descritiva - Bitcoin

Atividade 1: SME0620 - Estatística I

São Carlos
2024

Índice

1. Introdução.....	2
2. Metodologia.....	3
2.1. Base de dados.....	3
2.2. Classificação das variáveis.....	12
3. Resultados.....	12
3.1. Medidas de posição.....	13
3.2. Medidas de dispersão.....	15
3.3. Tabelas de frequência.....	16
3.4. Representações gráficas.....	20
3.5. Diagrama de dispersão - Quantitativas.....	22
3.6. Comparação quantitativa e qualitativa.....	26
4. Conclusão.....	27

1. Introdução

Desde que surgiu em 2009, o Bitcoin provocou uma verdadeira revolução no cenário financeiro global, introduzindo o conceito de criptomoedas como alternativas digitais às moedas tradicionais, operando fora do controle de instituições centralizadas. Utilizando criptografia para assegurar transações e controlar a emissão de novas unidades, o Bitcoin se destaca pelo uso da blockchain, uma tecnologia que promove transparência e resiste a interferências externas.

A mineração de Bitcoin é o processo pelo qual novos bitcoins são gerados. Esse processo exige um considerável poder computacional para resolver problemas matemáticos complexos, essencial para validar transações na rede e manter sua segurança. Essa atividade é motivada pelas recompensas em bloco, que são reduzidas pela metade periodicamente nos eventos de halving, aproximadamente a cada quatro anos. Esses eventos são cruciais, pois reduzem a oferta de novos bitcoins, o que teoricamente pode elevar seu preço se a demanda se mantiver estável ou aumentar.

A trajetória de preços do Bitcoin, disponível publicamente de 2010 a 2024, fornece um vasto campo para análise estatística, oferecendo insights sobre o efeito dos halvings no mercado. Além disso, este estudo incluirá o Índice VIX, conhecido como o "Índice do medo", que mede a volatilidade esperada baseada nas opções do índice S&P 500, de forma a analisar uma correlação entre este indicador e o preço da criptomoeda.

Sendo uma criptomoeda, seu mercado é extremamente volátil e suscetível à especulação, e estes diversos fatores de influência criam uma dificuldade grande para o investidor médio investir neste ativo ainda que confie na análise fundamentalista do projeto. Assim sendo, pretendemos responder a seguinte pergunta por meio desta análise estatística: Qual a melhor estratégia de investimento no Bitcoin baseado no tempo desde o último halving e o sentimento geral do mercado, tendo em vista um retorno a curto e longo prazo satisfatórios?

Portanto, não só este trabalho tem como objetivo elucidar o impacto dos halvings nos preços do Bitcoin, como também explorar como indicadores de mercados financeiros mais amplos, como o VIX, podem ser usados para prever tendências desse ativo digital, oferecendo valiosas perspectivas que poderiam beneficiar as estratégias de investimento em criptomoedas.

2. Metodologia

2.1. Base de dados

O histórico de preço do Bitcoin está disponível amplamente na internet, com preços registrados desde 2010. Optamos por organizar o histórico de preços de mês a mês, já que o preço médio de 30 dias é um bom representativo do preço enquanto diminui os efeitos da volatilidade diária deste mercado, criando um total de 167 linhas na base de dados tornando viável a pesquisa estatística.

Os preços mês a mês foram retirados de duas fontes: Do dia 01/08/2010 até o dia 01/10/2014, o histórico foi baixado no site Investing.com, e de 01/10/2014 até 01/05/2024 os preços foram retirados do site Yahoo Finanças. Como o Yahoo Finanças não disponibiliza o preço de fechamento do mês como o investing.com fez, foi necessário considerar o preço de fechamento como o preço de abertura do mês seguinte, uma aproximação com pouca variação (Volatilidade de um dia de mercado).

Ambas plataformas financeiras são populares pois oferecem uma ampla gama de informações sobre mercados financeiros. Investing.com fornece cotações em tempo real, gráficos, notícias, análises, dados técnicos e ferramentas financeiras, sendo uma fonte essencial para investidores interessados em ações, commodities, moedas, índices e futuros. Yahoo Finanças, por sua vez, oferece cotações de ações, notícias, gráficos interativos, relatórios de pesquisa e ferramentas de gestão de portfólio, destacando-se por sua interface amigável e facilidade de navegação.

Além disso, o valor do indicador VIX para o período de 01/08/2010 até 01/05/2024 foi retirado do site CBOE (Chicago Board Options Exchange), que é uma das maiores bolsas de opções do mundo. Porém, a sua plataforma disponibiliza apenas o valor do indicador de dia a dia, tornando necessário o cálculo do valor médio do indicador durante o mês. Optamos por realizar este cálculo utilizando o valor de fechamento de cada dia, pois este representa de forma mais precisa o sentimento do mercado durante o pregão.

Apesar das fontes de dados serem bem organizadas, foram necessários dois cálculos extras na planilha para tornar possíveis as análises que pretendíamos obter: a coluna Months_Since_Having e a de Fechamento. A primeira contabiliza a quantidade de meses que se passaram desde o último halving do bitcoin, e a segunda demonstra se o preço de fechamento do mês foi maior ou menor do que seu preço de abertura.

Toda a manipulação para junção dos dados das três diferentes fontes e a formatação dos mesmos para permitir os cálculos por python foi realizada pela ferramenta do Google Sheets. Também é importante deixar claro que todos os valores de preço da base de dados estão em Dólar.

Tabela 1: Base de dados utilizada na análise

Date	Months Since Having	Open	High	Low	Close	Fecha mento	VIX
01/05/2024	1	60 609,50	60 780,50	56 555,29	59 123,43	Queda	14,52
01/04/2024	0	71 333,48	72 715,36	59 120,07	60 636,86	Queda	16,14
01/03/2024	47	61 168,06	73 750,07	59 323,91	71 333,65	Alta	13,79
01/02/2024	46	42 569,76	63 913,13	41 879,19	61 198,38	Alta	13,98
01/01/2024	45	42 280,23	48 969,37	38 521,89	42 582,61	Alta	13,39
01/12/2023	44	37 718,01	44 705,52	37 629,36	42 265,19	Alta	12,72
01/11/2023	43	34 657,27	38 415,34	34 133,44	37 712,75	Alta	14,02
01/10/2023	42	26 967,40	35 150,43	26 558,32	34 667,78	Alta	18,89
01/09/2023	41	25 934,02	27 488,76	24 930,30	26 967,92	Alta	15,17
01/08/2023	40	29 230,87	30 176,80	25 409,11	25 931,47	Queda	15,85
01/07/2023	39	30 471,85	31 814,52	28 934,29	29 230,11	Queda	13,93
01/06/2023	38	27 218,41	31 389,54	24 797,17	30 477,25	Alta	14,00
01/05/2023	37	29 227,10	29 820,13	25 878,43	27 219,66	Queda	17,64
01/04/2023	36	28 473,33	31 005,61	27 070,85	29 268,81	Alta	17,82
01/03/2023	35	23 150,93	29 159,90	19 628,25	28 478,48	Alta	21,64
01/02/2023	34	23 137,84	25 134,12	21 460,09	23 147,35	Alta	20,12
01/01/2023	33	16 547,91	23 919,89	16 521,23	23 139,28	Alta	20,17
01/12/2022	32	17 168,00	18 318,53	16 398,14	16 547,50	Queda	21,78

01/11/2022	31	20 494,90	21 446,89	15 599,05	17 168,57	Queda	23,30
01/10/2022	30	19 431,11	20 988,39	18 319,82	20 495,77	Alta	30,01
01/09/2022	29	20 050,50	22 673,82	18 290,31	19 431,79	Queda	27,34
01/08/2022	28	23 336,72	25 135,59	19 600,79	20 049,76	Queda	22,17
01/07/2022	27	19 820,47	24 572,58	18 966,95	23 336,90	Alta	25,00
01/06/2022	26	31 792,55	31 957,29	17 708,62	19 784,73	Queda	28,23
01/05/2022	25	37 713,27	39 902,95	26 350,49	31 792,31	Queda	29,31
01/04/2022	24	45 554,16	47 313,48	37 697,94	37 714,88	Queda	24,37
01/03/2022	23	43 194,50	48 086,84	37 260,20	45 538,68	Alta	26,97
01/02/2022	22	38 481,77	45 661,17	34 459,22	43 193,23	Alta	25,75
01/01/2022	21	46 311,75	47 881,41	33 184,06	38 483,13	Queda	23,18
01/12/2021	20	56 907,96	59 041,68	42 874,62	46 306,45	Queda	21,35
01/11/2021	19	61 320,45	68 789,63	53 569,77	57 005,43	Queda	18,50
01/10/2021	18	43 816,74	66 930,39	43 320,02	61 318,96	Alta	17,87
01/09/2021	17	47 099,77	52 853,77	39 787,61	43 790,89	Queda	19,82
01/08/2021	16	41 460,84	50 482,08	37 458,00	47 166,69	Alta	17,47
01/07/2021	15	35 035,98	42 235,55	29 360,96	41 626,20	Alta	17,60
01/06/2021	14	37 293,79	41 295,27	28 893,62	35 040,84	Queda	16,96
01/05/2021	13	57 714,66	59 519,36	30 681,50	37 332,86	Queda	19,76
01/04/2021	12	58 926,56	64 863,10	47 159,48	57 750,18	Queda	17,42
01/03/2021	11	45 159,50	61 683,86	45 115,09	58 918,83	Alta	21,84

01/02/2021	10	33 114,58	58 330,57	32 384,23	45 137,77	Alta	23,14
01/01/2021	9	28 994,01	41 946,74	28 722,76	33 114,36	Alta	24,91
01/12/2020	8	19 633,77	29 244,88	17 619,53	29 001,72	Alta	22,37
01/11/2020	7	13 781,00	19 749,26	13 243,16	19 625,84	Alta	25,00
01/10/2020	6	10 795,25	14 028,21	10 416,69	13 781,00	Alta	29,44
01/09/2020	5	11 679,32	12 067,08	9 916,49	10 784,49	Queda	27,65
01/08/2020	4	11 322,57	12 359,06	11 012,42	11 680,82	Alta	22,89
01/07/2020	3	9 145,99	11 415,86	8 977,02	11 323,47	Alta	26,84
01/06/2020	2	9 463,61	10 199,57	8 975,53	9 137,99	Queda	31,12
01/05/2020	1	8 672,78	9 996,74	8 374,32	9 461,06	Alta	30,90
01/04/2020	0	6 437,32	9 440,65	6 202,37	8 658,55	Alta	41,45
01/03/2020	44	8 599,76	9 167,70	4 106,98	6 438,64	Queda	57,74
01/02/2020	43	9 346,36	10 457,63	8 492,93	8 599,51	Queda	19,63
01/01/2020	42	7 194,89	9 553,13	6 915,00	9 350,53	Alta	13,94
01/12/2019	41	7 571,62	7 743,43	6 540,05	7 193,60	Queda	13,76
01/11/2019	40	9 193,99	9 505,05	6 617,17	7 569,63	Queda	12,52
01/10/2019	39	8 299,72	10 021,74	7 446,99	9 199,58	Alta	15,47
01/09/2019	38	9 630,59	10 898,76	7 830,76	8 293,87	Queda	15,56
01/08/2019	37	10 077,44	12 273,82	9 421,63	9 630,66	Queda	18,98
01/07/2019	36	10 796,93	13 129,53	9 163,13	10 085,63	Queda	13,31
01/06/2019	35	8 573,84	13 796,49	7 564,49	10 817,16	Alta	15,84

01/05/2019	34	5 350,91	9 008,31	5 347,65	8 574,50	Alta	16,72
01/04/2019	33	4 105,36	5 642,04	4 096,90	5 350,73	Alta	12,95
01/03/2019	32	3 853,76	4 296,81	3 733,75	4 105,40	Alta	14,49
01/02/2019	31	3 460,55	4 210,64	3 391,02	3 854,79	Alta	15,23
01/01/2019	30	3 746,71	4 109,02	3 400,82	3 457,79	Queda	19,57
01/12/2018	29	4 024,46	4 309,38	3 191,30	3 742,70	Queda	24,95
01/11/2018	28	6 318,14	6 552,16	3 585,06	4 017,27	Queda	19,39
01/10/2018	27	6 619,85	6 965,06	6 236,47	6 317,61	Queda	19,35
01/09/2018	26	7 044,81	7 388,43	6 197,52	6 625,56	Queda	12,91
01/08/2018	25	7 769,04	7 769,04	5 971,05	7 037,58	Queda	12,55
01/07/2018	24	6 411,68	8 424,27	6 136,42	7 780,44	Alta	13,15
01/06/2018	23	7 500,70	7 754,89	5 826,41	6 404,00	Queda	13,68
01/05/2018	22	9 251,47	9 964,50	7 090,68	7 494,17	Queda	14,12
01/04/2018	21	7 003,06	9 745,32	6 526,87	9 240,55	Alta	18,27
01/03/2018	20	10 385,00	11 704,10	6 683,93	6 973,53	Queda	19,02
01/02/2018	19	10 237,30	11 958,50	6 048,26	10 397,90	Alta	22,46
01/01/2018	18	14 112,20	17 712,40	9 402,29	10 221,10	Queda	11,06
01/12/2017	17	10 198,60	20 089,00	9 694,65	14 156,40	Alta	10,26
01/11/2017	16	6 440,97	11 517,40	5 519,01	10 233,60	Alta	10,54
01/10/2017	15	4 341,05	6 470,43	4 164,05	6 468,40	Alta	10,13
01/09/2017	14	4 701,76	4 975,04	2 946,62	4 338,71	Queda	10,44

01/08/2017	13	2 871,30	4 736,05	2 668,59	4 703,39	Alta	11,98
01/07/2017	12	2 492,60	2 916,14	1 843,03	2 875,34	Alta	10,26
01/06/2017	11	2 288,33	2 999,91	2 212,96	2 480,84	Alta	10,51
01/05/2017	10	1 348,30	2 763,71	1 348,30	2 286,41	Alta	10,86
01/04/2017	9	1 071,71	1 347,91	1 061,09	1 347,89	Alta	13,14
01/03/2017	8	1 180,04	1 280,31	903,71	1 071,79	Queda	11,90
01/02/2017	7	970,94	1 200,39	946,69	1 179,97	Alta	11,53
01/01/2017	6	963,66	1 191,10	755,76	970,40	Alta	11,61
01/12/2016	5	746,05	979,40	746,05	963,74	Alta	12,47
01/11/2016	4	701,34	756,24	678,16	745,69	Alta	15,24
01/10/2016	3	609,93	720,40	609,48	700,97	Alta	14,59
01/09/2016	2	575,55	628,82	570,81	609,73	Alta	14,22
01/08/2016	1	624,60	626,12	531,33	575,47	Queda	12,40
01/07/2016	0	672,52	704,97	611,83	624,68	Queda	13,16
01/06/2016	45	531,11	777,99	525,64	673,34	Alta	17,77
01/05/2016	44	448,48	553,96	437,39	531,39	Alta	14,85
01/04/2016	43	416,76	467,96	415,83	448,32	Alta	14,30
01/03/2016	42	437,92	439,65	394,04	416,73	Queda	15,85
01/02/2016	41	369,35	448,05	367,96	437,70	Alta	22,52
01/01/2016	40	430,72	462,93	354,91	368,77	Queda	23,72
01/12/2015	39	377,41	469,10	349,46	430,57	Alta	18,03

01/11/2015	38	315,01	495,56	301,00	377,32	Alta	16,21
01/10/2015	37	236,00	334,17	235,62	314,17	Alta	16,79
01/09/2015	36	230,26	259,18	225,12	236,06	Alta	24,38
01/08/2015	35	284,69	285,71	199,57	230,06	Queda	19,43
01/07/2015	34	263,35	314,39	253,51	284,65	Alta	14,35
01/06/2015	33	230,23	267,87	221,30	263,07	Alta	14,34
01/05/2015	32	235,94	247,80	228,57	230,19	Queda	13,34
01/04/2015	31	244,22	261,80	214,87	236,15	Queda	13,49
01/03/2015	30	254,28	300,04	236,51	244,22	Queda	14,81
01/02/2015	29	216,87	265,61	212,01	254,26	Alta	15,90
01/01/2015	28	320,43	320,43	171,51	217,46	Queda	19,12
01/12/2014	27	378,25	384,04	304,23	320,19	Queda	16,29
01/11/2014	26	338,65	457,09	320,63	378,05	Alta	13,41
01/10/2014	25	388,20	407,70	294,90	338,65	Queda	18,06
01/09/2014	24	481,80	498,50	372,60	388,20	Queda	13,47
01/08/2014	23	589,50	608,20	450,00	481,80	Queda	13,49
01/07/2014	22	635,10	652,50	564,00	589,50	Queda	12,30
01/06/2014	21	627,90	676,50	560,80	635,10	Alta	11,54
01/05/2014	20	445,60	629,00	420,20	627,90	Alta	12,48
01/04/2014	19	444,70	549,00	355,70	445,60	Alta	14,20
01/03/2014	18	573,90	695,40	262,80	444,70	Queda	14,84

01/02/2014	17	938,80	969,20	91,70	573,90	Queda	15,47
01/01/2014	16	805,90	1 093,40	771,00	938,80	Alta	14,24
01/12/2013	15	1 205,70	1 239,90	454,90	805,90	Queda	14,19
01/11/2013	14	211,20	1 241,90	209,70	1 205,70	Alta	12,92
01/10/2013	13	141,90	233,40	109,70	211,20	Alta	15,41
01/09/2013	12	141,00	148,90	121,30	141,90	Alta	14,69
01/08/2013	11	106,20	148,70	101,00	141,00	Alta	14,21
01/07/2013	10	97,50	111,70	65,40	106,20	Alta	13,97
01/06/2013	9	128,80	130,10	88,00	97,50	Queda	17,27
01/05/2013	8	139,20	140,10	79,00	128,80	Queda	13,49
01/04/2013	7	93,00	266,00	50,00	139,20	Alta	13,97
01/03/2013	6	33,40	95,70	32,90	93,00	Alta	13,03
01/02/2013	5	20,40	34,50	18,00	33,40	Alta	14,07
01/01/2013	4	13,50	21,40	13,20	20,40	Alta	13,51
01/12/2012	3	12,60	13,90	12,40	13,50	Alta	17,34
01/11/2012	2	11,20	12,60	10,30	12,60	Alta	16,70
01/10/2012	1	12,40	13,10	9,70	11,20	Queda	16,28
01/09/2012	0	10,20	12,70	9,70	12,40	Alta	15,28
01/08/2012	-1	9,40	15,40	7,60	10,20	Alta	15,69
01/07/2012	-2	6,70	9,70	6,40	9,40	Alta	17,56
01/06/2012	-3	5,20	6,80	5,20	6,70	Alta	21,13

01/05/2012	-4	4,90	5,20	4,80	5,20	Alta	21,00
01/04/2012	-5	4,90	5,50	4,60	4,90	Alta	17,82
01/03/2012	-6	4,90	5,40	4,30	4,90	Alta	16,17
01/02/2012	-7	5,50	6,20	3,90	4,90	Queda	18,42
01/01/2012	-8	4,70	7,20	4,60	5,50	Alta	20,23
01/12/2011	-9	3,00	5,00	2,60	4,70	Alta	25,05
01/11/2011	-10	3,30	3,40	2,00	3,00	Queda	31,94
01/10/2011	-11	5,10	5,30	2,00	3,30	Queda	32,83
01/09/2011	-12	8,20	8,70	4,20	5,10	Queda	36,53
01/08/2011	-13	13,40	13,60	5,70	8,20	Queda	35,03
01/07/2011	-14	16,10	16,70	11,00	13,40	Queda	19,23
01/06/2011	-15	8,70	31,90	8,30	16,10	Alta	19,15
01/05/2011	-16	3,50	9,50	2,50	8,70	Alta	16,91
01/04/2011	-17	0,80	4,20	0,60	3,50	Alta	16,24
01/03/2011	-18	0,90	1,00	0,70	0,80	Queda	20,72
01/02/2011	-19	0,50	1,10	0,50	0,90	Alta	17,43
01/01/2011	-20	0,30	0,90	0,30	0,50	Alta	17,32
01/12/2010	-21	0,20	0,30	0,20	0,30	Alta	17,57
01/11/2010	-22	0,20	0,50	0,10	0,20	Alta	20,10
01/10/2010	-23	0,10	0,20	0,00	0,20	Alta	20,37
01/09/2010	-24	0,10	0,20	0,10	0,10	Alta	22,52

01/08/2010	-25	0,10	0,10	0,00	0,10	Alta	24,75
------------	-----	------	------	------	------	------	-------

2.2. Classificação das variáveis

- **Date:** Variável qualitativa ordinal, pois representa datas que possuem uma ordem cronológica.
- **Months_Since_Halving:** Variável quantitativa discreta, pois é uma contagem do número de meses.
- **Open:** Variável quantitativa contínua, pois representa preços de abertura que podem assumir qualquer valor dentro de um intervalo.
- **High:** Variável quantitativa contínua, pois representa o preço máximo que pode assumir qualquer valor dentro de um intervalo.
- **Low:** Variável quantitativa contínua, pois representa o preço mínimo que pode assumir qualquer valor dentro de um intervalo.
- **Close:** Variável quantitativa contínua, pois representa o preço de fechamento que pode assumir qualquer valor dentro de um intervalo.
- **Fechamento:** Variável qualitativa nominal, pois o indicador de alta ou queda não tem ordem ou hierarquia natural.
- **VIX:** Variável quantitativa contínua, pois representa um índice que pode assumir qualquer valor dentro de um intervalo.

3. Resultados

Script 1: Importando bibliotecas e definindo o Dataframe no Colab:

```
#Installs
!pip install openpyxl matplotlib seaborn
!pip install openpyxl

#Imports
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from IPython.display import display

# Carregando a planilha
df = pd.read_excel('BTC_BD.xlsx')

# Exibir as primeiras linhas para verificar o carregamento correto
print(df.head())
```

3.1. Medidas de posição

Tabela 1 - Moda, média e quartis (1, 2 e 3) das variáveis de interesse.

Campo	Moda	Média	Q1	Mediana(Q2)	Q3
Open	0.1	11361.94	230.24	2390.47	15938.99
High	0.2	13357.28	265.71	2958.02	20004.07
Low	0	9835	202.1	2027.99	15010.08
Close	4.9	11720.73	231.66	2678.09	17013.3
Fechamento	Alta	-	-	-	-
VIX	10.2645	18.22	13.97	16.71	20.64

A análise descritiva dos preços históricos do Bitcoin (2010-2024) revela uma grande dispersão nos valores. As modas dos preços de abertura, máxima, mínima e fechamento são baixas, refletindo os valores frequentes no início da série. As médias dos preços são elevadas, indicando influência de picos recentes. Os quartis (Q1, mediana e Q3) mostram que a maioria dos valores está abaixo desses picos, indicando a presença de valores extremamente altos em momentos específicos. O índice VIX, com uma média de 18.22 e mediana de 16.71, aponta para uma volatilidade significativa no mercado, sugerindo correlações potenciais com os preços do Bitcoin.

Script 2: Cálculo das medidas de posição:

```
# Função para calcular as medidas de posição, ignorando médias e medianas para variáveis qualitativas

def calcular_medidas(df, colunas_qualitativas):

    medidas = []

    for coluna in df.columns:

        if coluna in colunas_qualitativas:

            moda = df[coluna].mode().iloc[0]

            medidas.append([coluna, moda, '-', '-', '-', '-'])

        else:

            moda = df[coluna].mode().iloc[0]
```

```

        media = df[coluna].mean().round(2)

        q1 = df[coluna].quantile(0.25).round(2)

        mediana = df[coluna].median().round(2)

        q3 = df[coluna].quantile(0.75).round(2)

        medidas.append([coluna, moda, media, q1, mediana, q3])

    return pd.DataFrame(medidas, columns=['Campo', 'Moda', 'Média',
'Q1', 'Mediana (Q2)', 'Q3'])

# Selecionar as colunas de interesse

colunas_interesse = ['Open', 'High', 'Low', 'Close', 'Fechamento',
'VIX']

colunas_qualitativas = ['Fechamento']

# Calcular e exibir as medidas

medidas = calcular_medidas(df[colunas_interesse],
colunas_qualitativas)

# Exibir a tabela de medidas

display(medidas)

```

3.2. Medidas de dispersão

Tabela 2 - Amplitude, intervalo interquartil, variância, desvio padrão e coeficiente de variação das variáveis de interesse.

Campo	Amplitude	Intervalo Interquartil (IQR)	Variância	Desvio Padrão	Coeficiente de Variação (%)
Open	71333.38	15708.75	279602440.34	16721.32	147.17
High	73749.97	19738.36	373644515.04	19329.89	144.71
Low	59323.91	14807.97	210179140.53	14497.56	147.41
Close	71333.55	16781.64	292708340.02	17108.72	145.97
VIX	47.61	6.67	41.61	6.45	35.4

A análise das medidas de dispersão dos preços históricos do Bitcoin revela uma significativa variabilidade. As amplitudes para os preços de abertura, máxima, mínima e fechamento estão acima de 59.000 dólares, destacando grandes flutuações. O intervalo interquartil (IQR) para esses preços mostra uma dispersão considerável no mercado. As variâncias e os desvios padrão elevados, especialmente para os preços de fechamento, reforçam a alta volatilidade. O coeficiente de variação (CV) para os preços do Bitcoin ultrapassa 144%, indicando alta variabilidade relativa, enquanto o VIX apresenta menor dispersão (CV de 35,4%), refletindo uma volatilidade esperada mais estável no mercado financeiro.

Script 3: Cálculo das medidas de dispersão:

```
# Calcular as medidas de dispersão

def calcular_medidas_dispersao(df):

    dispersao = pd.DataFrame()

    dispersao['Amplitude'] = (df.max() - df.min()).round(2)

    dispersao['Intervalo Interquartil (IQR)'] = (df.quantile(0.75) -
df.quantile(0.25)).round(2)

    dispersao['Variância'] = df.var().round(2)

    dispersao['Desvio Padrão'] = df.std().round(2)

    dispersao['Coeficiente de Variação (%)'] = (df.std() / df.mean()
* 100).round(2)
```



```

        return pd.DataFrame(dispersao, columns=['Amplitude', 'Intervalo
Interquartil (IQR)', 'Variância', 'Desvio Padrão', 'Coeficiente de
Variação (%)'])

# Selecionar as colunas de interesse

colunas_interesse = ['Open', 'High', 'Low', 'Close', 'VIX']

# Calcular e exibir as medidas de dispersão

medidas_dispersao = calcular_medidas_dispersao(df[colunas_interesse])

# Renomear o índice para "Campo"

medidas_dispersao.index.name = "Campo"

# Exibir a tabela de medidas de dispersão

display(medidas_dispersao)

```

3.3. Tabelas de frequência

Tabela 3 - Frequência da variável qualitativa (Fechamento).

Fechamento	Frequência
Alta	99
Queda	67

A Tabela 3 mostra que o preço de fechamento do Bitcoin teve alta em 99 meses e queda em 67 meses, indicando uma ligeira predominância de altas (aproximadamente 60%) ao longo do período estudado.

Tabela 4 - Frequência da variável quantitativa (Close)

Close (em intervalo)	Frequência
(-71.234, 7133.455]	99
(7133.455, 14266.81]	24
(28533.519, 35666.874]	8
(14266.81, 21400.165]	7
(21400.165, 28533.519]	7
(35666.874, 42800.229]	7
(42800.229, 49933.584]	6
(57066.939, 64200.294]	6
(49933.584, 57066.939]	1
(64200.294, 71333.648]	1

A Tabela 4 revela que a maioria dos preços de fechamento (99 meses) está no intervalo até 7133.455 dólares, com frequências decrescentes nos intervalos superiores. Isso indica que os preços do Bitcoin foram majoritariamente baixos, com menos meses registrando valores acima de 50.000 dólares, refletindo a volatilidade e a trajetória de crescimento ao longo do tempo.

Script 4: Cálculo das medidas de frequência:

```
# Renomear categorias 'Fechamento'

df['Fechamento'] = df['Fechamento'].replace({0: 'Queda', 1: 'Alta'})

# Calcular a frequência das categorias da variável qualitativa
'Fechamento'

freq_qualitative = df['Fechamento'].value_counts().reset_index()

freq_qualitative.columns = ['Categoria', 'Frequência']

freq_qualitative = freq_qualitative.set_index('Categoria')

# Exibir a tabela de frequência da variável qualitativa

print("Frequência da variável qualitativa (Fechamento):")
```

```

display(pd.DataFrame(freq_qualitative))

# Gráfico de barras para a variável qualitativa
ax1 = freq_qualitative.plot(kind='bar', legend=False)

plt.title('Frequência de Fechamentos de Bitcoin')

plt.xlabel('Categoria de Fechamento')

plt.ylabel('Frequência')

# Adicionar rótulos nas colunas
for p in ax1.patches:
    ax1.annotate(str(p.get_height()), (p.get_x() * 1.005,
p.get_height() * 1.005))

plt.show()

# Calcular a frequência dos valores da variável quantitativa 'Close'
freq_quantitative = pd.cut(df['Close'],
bins=10).value_counts().reset_index()

freq_quantitative.columns = ['Intervalo de Fechamento',
'Frequência']

freq_quantitative = freq_quantitative.set_index('Intervalo de
Fechamento')

# Exibir a tabela de frequência da variável quantitativa
print("\nFrequência da variável quantitativa (Close):")

display(pd.DataFrame(freq_quantitative))

```

```
# Histograma para a variável quantitativa

ax2 = df['Close'].plot(kind='hist', bins=10)

plt.title('Distribuição dos Fechamentos de Bitcoin')

plt.xlabel('Preço de Fechamento')

plt.ylabel('Frequência')


# Adicionar rótulos nas colunas

n, bins, patches = ax2.hist(df['Close'], bins=10)

for i in range(len(patches)):

    ax2.annotate(str(int(n[i])), (patches[i].get_x() +
patches[i].get_width() / 2, n[i]),

                ha='center', va='bottom')

plt.show()
```

3.4. Representações gráficas

Figura 1 - Histogramas das variáveis quantitativas e gráfico de barras da variável qualitativa.

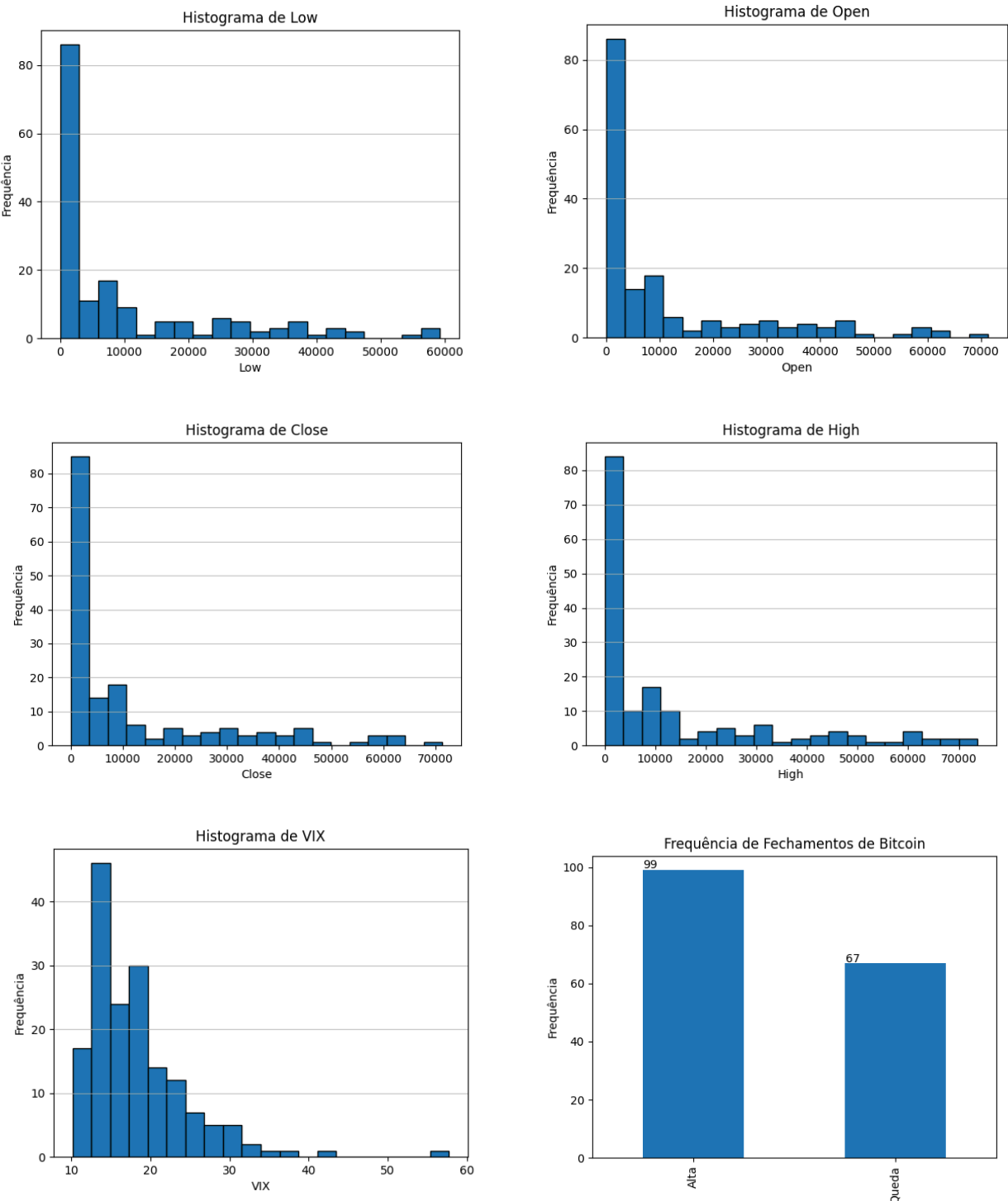
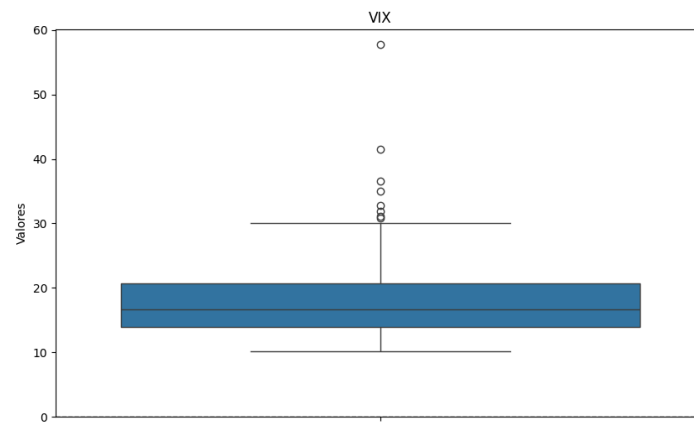
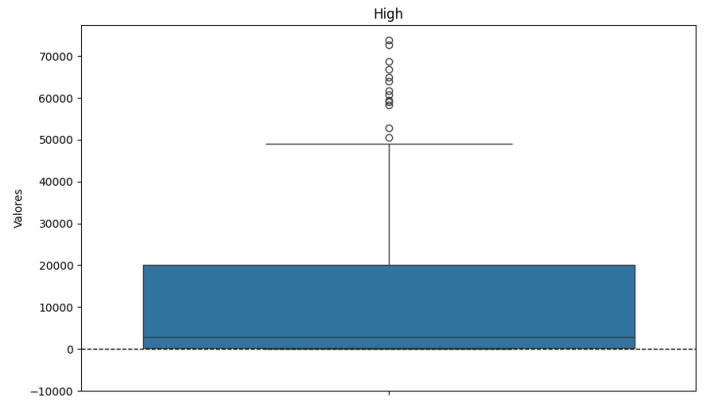
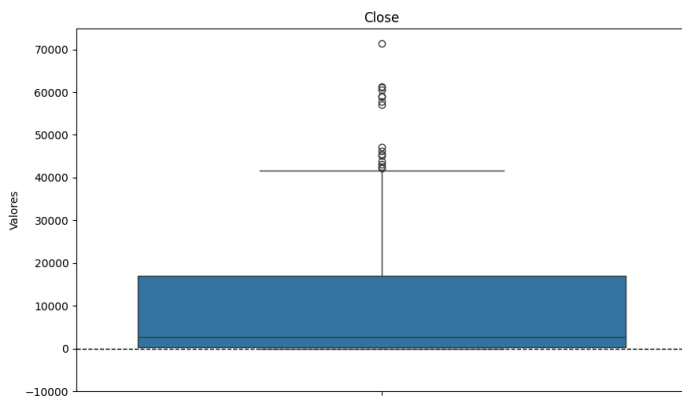
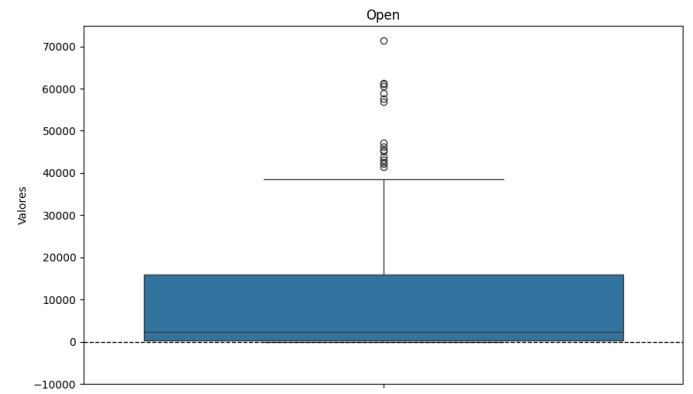
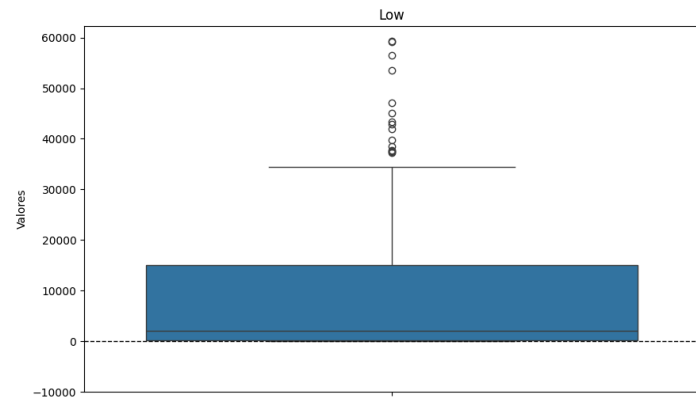


Figura 2 - Boxplot das variáveis qualitativas



3.5. Diagrama de dispersão - Quantitativas

Como o Bitcoin sofreu 4 halvings desde a sua concepção, tendo o mais novo ocorrido em abril de 2024, existem 3 períodos entre halvings a serem analisados na nossa base de dados:

- 1º período: entre 01/09/2012 e 01/07/2016;
- 2º período: entre 01/07/2016 e 01/04/2020;
- 3º período: entre 01/04/2020 e 01/04/2024.

Assim sendo, consideramos mais válido para esta análise de relação entre as variáveis quantitativas Close e Months Since Halving que fossem calculados a covariância amostral e o coeficiente de Pearson para cada período para no final ser realizada a média dos resultados, assim diminuindo os efeitos da especulação na base de dados e trazendo números finais mais limpos. A mesma lógica foi aplicada para os diagramas de dispersão, para facilitar a visualização de padrões.

Figura 3 - Diagrama de dispersão referente ao período 1 entre Close e Months Since Halving

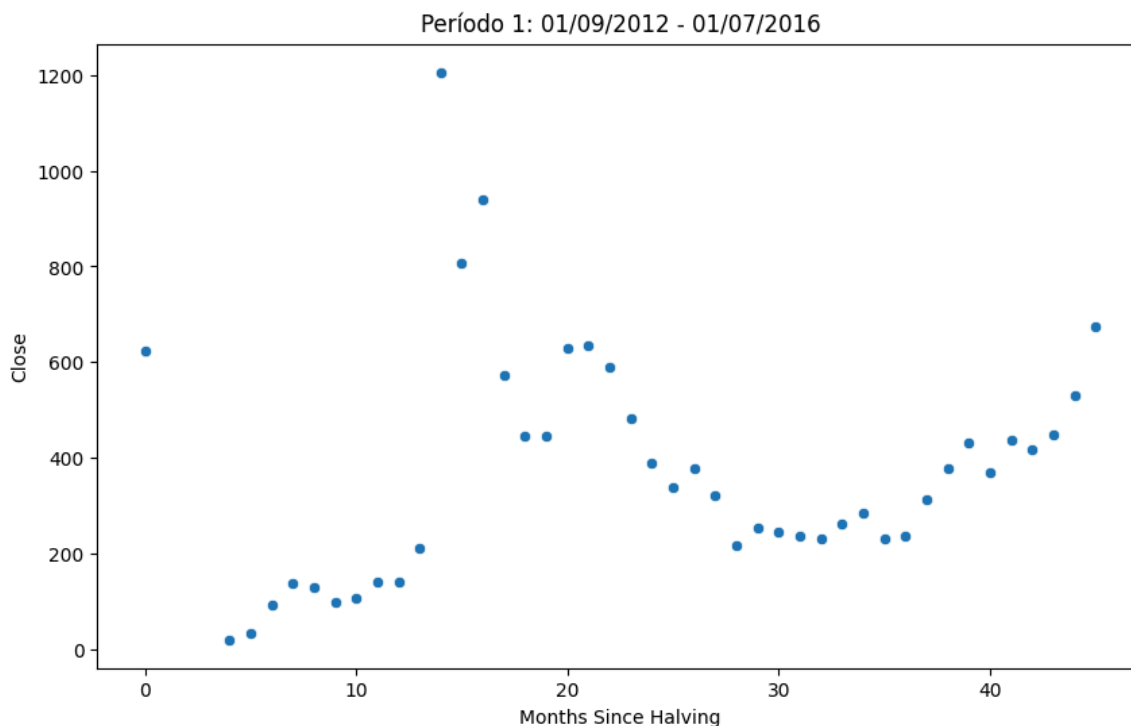


Figura 4 - Diagrama de dispersão referente ao período 2 entre Close e Months Since Halving

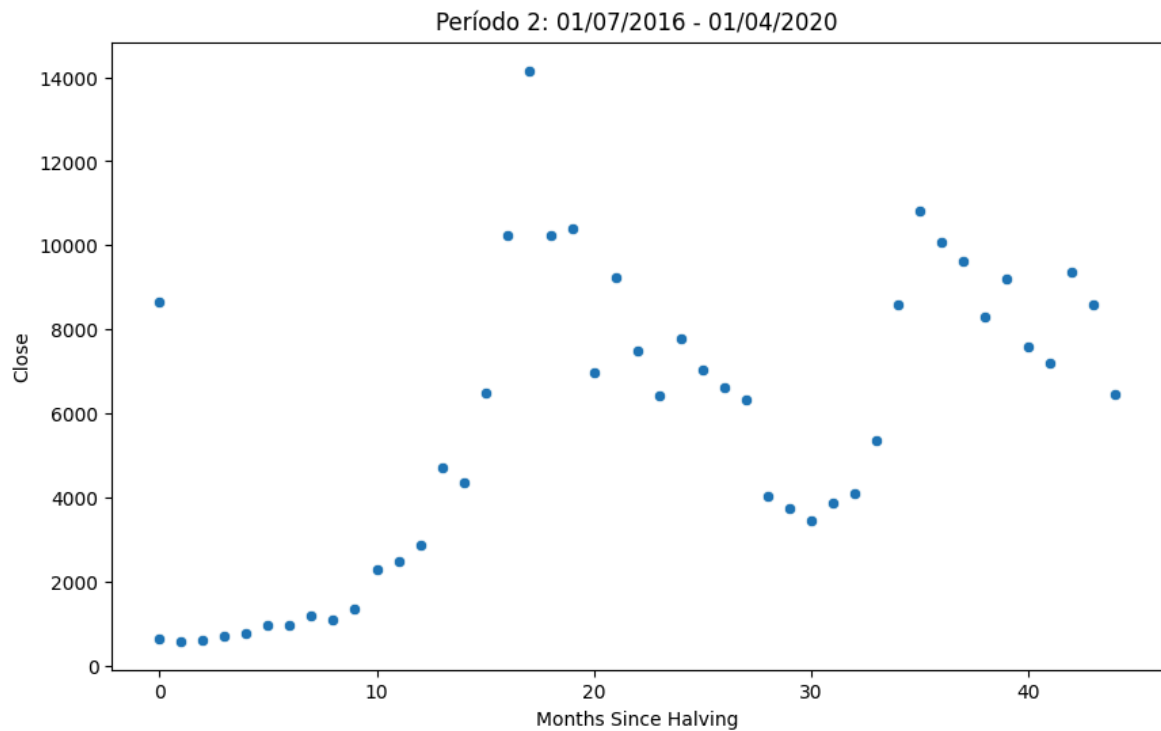


Figura 5 - Diagrama de dispersão referente ao período 3 entre Close e Months Since Halving

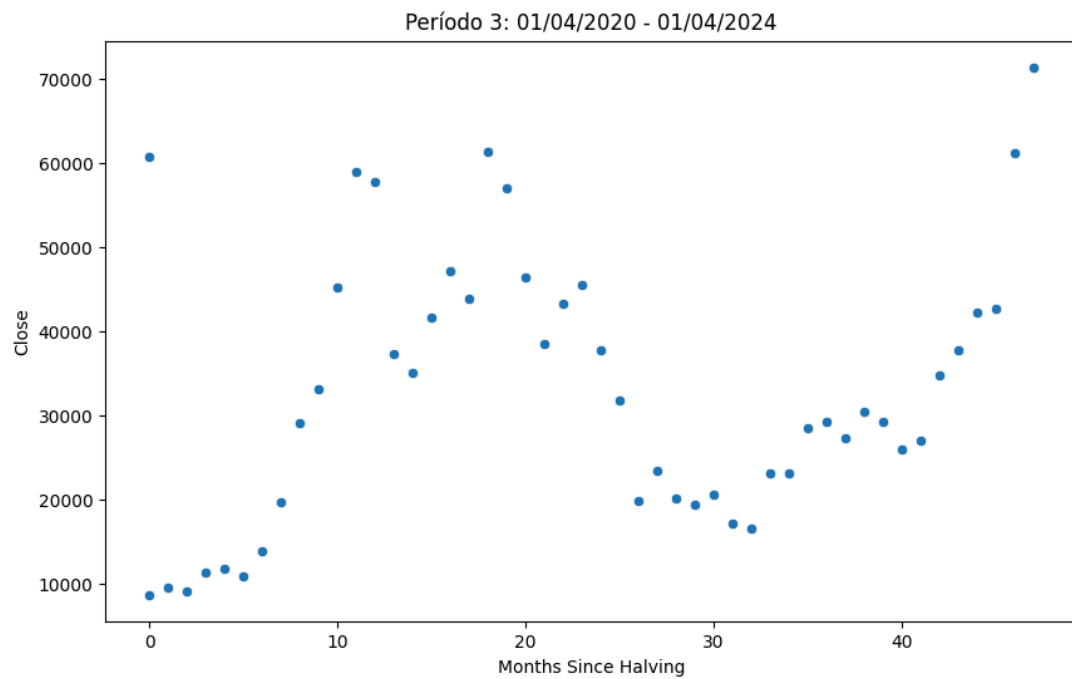


Tabela 5 - Covariância amostral e coeficiente de Pearson entre a Close e Months Since Halving

Period	Covariance	Pearson Correlation (r)
01/09/2012 - 01/07/2016	429.87	0.14
01/07/2016 - 01/04/2020	28215.62	0.59
01/04/2020 - 01/04/2024	46755.72	0.2
Average	25133.73	0.31

Assim, por meio da covariância amostral média de 25133.73 entre os três períodos, podemos inferir a partir do sinal positivo que, conforme o tempo desde o último halving passa, o preço do Bitcoin tende a subir. No entanto, a magnitude da covariância depende das unidades das variáveis, o que pode tornar difícil a interpretação direta sem contexto adicional.

Além do mais, a partir do coeficiente de Pearson médio de 0.31, é possível determinar que as duas variáveis qualitativas possuem uma correlação mediana, provavelmente impactada pela alta complexidade do mercado de criptomoedas, porém a sua análise não pode ser desconsiderada para melhora de estratégias de investimento.

Por último, os três diagramas de dispersão demonstram um claro padrão no preço do criptoativo nos meses seguintes após um halving. O preço inicia uma subida forte que atinge seu pico por volta do 16º mês, a partir do qual declina até atingir seu mínimo por volta do seu 30º mês e em seguida retorna a uma subida com seu pico no próximo halving.

Script 5: Cálculo da covariância e coeficiente de Pearson entre Close e Months Since Halving:

```
# Convertendo as colunas para o tipo correto
df['Close'] = pd.to_numeric(df['Close'], errors='coerce')
df['Months_Since_Halving'] = pd.to_numeric(df['Months_Since_Halving'],
errors='coerce')
df['Date'] = pd.to_datetime(df['Date'], errors='coerce')

# Filtrando os períodos de interesse
period1 = df[(df['Date'] >= '2012-09-01') & (df['Date'] <= '2016-07-01')]
period2 = df[(df['Date'] >= '2016-07-01') & (df['Date'] <= '2020-04-01')]
period3 = df[(df['Date'] >= '2020-04-01') & (df['Date'] <= '2024-04-01')]

# Função para calcular covariância e correlação
def calculate_cov_corr(df, x_col, y_col):
    covariance = np.cov(df[x_col], df[y_col])[0, 1].round(2)
    correlation = np.corrcoef(df[x_col], df[y_col])[0, 1].round(2)
    return covariance, correlation

# Calculando covariância e correlação para cada período
cov_corr_period1 = calculate_cov_corr(period1, 'Close',
'Months_Since_Halving')
cov_corr_period2 = calculate_cov_corr(period2, 'Close',
'Months_Since_Halving')
cov_corr_period3 = calculate_cov_corr(period3, 'Close',
'Months_Since_Halving')
```

```

# Criando a tabela de resultados
results = pd.DataFrame({
    'Period': ['01/09/2012 - 01/07/2016', '01/07/2016 - 01/04/2020',
'01/04/2020 - 01/04/2024'],
    'Covariance': [cov_corr_period1[0], cov_corr_period2[0],
cov_corr_period3[0]],
    'Pearson Correlation (r)': [cov_corr_period1[1], cov_corr_period2[1],
cov_corr_period3[1]]
})

# Calculando a média
mean_covariance = results['Covariance'].mean()
mean_correlation = results['Pearson Correlation (r)'].mean()

# Adicionando as médias à tabela
results.loc[len(results)] = ['Average', mean_covariance, mean_correlation]

# Exibindo a tabela
display(results)

# Plotando os diagramas de dispersão em figuras separadas

# Período 1
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Months_Since_Halving', y='Close', data=period1)
plt.title('Período 1: 01/09/2012 - 01/07/2016')
plt.xlabel('Months Since Halving')
plt.ylabel('Close')
plt.show()

# Período 2
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Months_Since_Halving', y='Close', data=period2)
plt.title('Período 2: 01/07/2016 - 01/04/2020')
plt.xlabel('Months Since Halving')
plt.ylabel('Close')
plt.show()

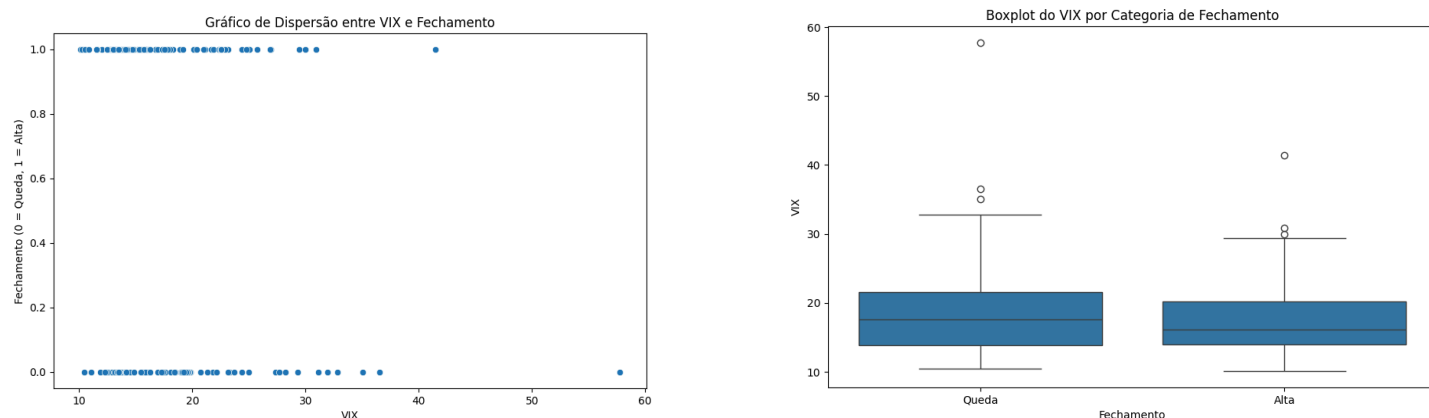
# Período 3
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Months_Since_Halving', y='Close', data=period3)
plt.title('Período 3: 01/04/2020 - 01/04/2024')
plt.xlabel('Months Since Halving')

```

```
plt.ylabel('Close')
plt.show()
```

3.6. Comparação quantitativa e qualitativa

Figura 6 - Boxplot e Dispersão para a relação Close x Fechamento.



Como pode ser analisado a partir do diagrama de dispersão das variáveis quantitativas VIX e Fechamento, não há um padrão claro que indique uma relação forte entre a volatilidade do mercado e o fechamento mensal do Bitcoin. O gráfico de dispersão mostra uma distribuição dispersa dos pontos de dados, sugerindo uma fraca correlação. O boxplot reforça essa observação ao demonstrar que as distribuições do VIX para os diferentes estados de fechamento ("Alta" e "Queda") se sobrepõem significativamente, indicando que a variação do VIX não está fortemente associada com o fechamento mensal do Bitcoin.

Esta conclusão é suportada pelos valores de covariância amostral e coeficiente de correlação linear de Pearson. A covariância amostral de -0.45618 indica uma relação inversa, sugerindo que quando o VIX aumenta, o fechamento do Bitcoin tende a ser uma queda. No entanto, a magnitude da covariância torna difícil a interpretação direta sem contexto adicional. O coeficiente de correlação de Pearson de aproximadamente -0.14371 confirma essa relação inversa, mas também indica que a correlação é muito fraca. Isso sugere que, apesar de uma ligeira tendência inversa, o VIX não é um preditor confiável para as variações no fechamento do Bitcoin, e outros fatores e métodos analíticos devem ser considerados para uma compreensão mais profunda das dinâmicas do mercado.

Script 6: Cálculo da covariância amostral e coeficiente de correlação de Pearson entre VIX e Fechamento:

```
# Converter a variável qualitativa 'Fechamento' em uma variável numérica
df['Fechamento_Num'] = df['Fechamento'].apply(lambda x: 1 if x == 'Alta'
else 0)
```

```
# Calcular a covariância amostral entre 'VIX' e 'Fechamento_Num'
covariance = df[['VIX', 'Fechamento_Num']].cov().iloc[0, 1]
print(f'Covariância amostral entre VIX e Fechamento: {covariance}')

# Calcular o coeficiente de correlação de Pearson entre 'VIX' e
'Fechamento_Num'
correlation = df['VIX'].corr(df['Fechamento_Num'])
print(f'Coeficiente de correlação entre VIX e Fechamento: {correlation}')
```

4. Conclusão

O objetivo inicial deste trabalho de pesquisa estatística era de responder a seguinte pergunta: Qual a melhor estratégia de investimento no Bitcoin baseado no tempo desde o último halving e o sentimento geral do mercado, tendo em vista um retorno a curto e longo prazo satisfatórios?

Sendo assim, a partir da base de dados que consideramos foi possível analisar que, ainda que a variação do índice VIX não tenha uma correlação forte com o preço do ativo, ele pode ser utilizado como uma ferramenta de análise do sentimento de volatilidade do mercado como um todo para uma tomada de decisões mais cautelosa.

Além do mais, existe uma clara relação entre o preço do Bitcoin e o tempo percorrido desde o último halving, havendo um padrão que se repetiu nos 3 períodos entre halvings de 2012 até 2024, algo que pode ser extremamente valioso para garantir um bom ponto de entrada e saída neste mercado volátil.

Assim, com base nesses fatos é possível criar uma estratégia a maximizar o retorno sobre o investimento no Bitcoin: Se o investidor comprar o ativo no mês do halving, vender este ativo 16 meses depois, recomprá-lo no 30º mês e mantê-lo em sua carteira para o mesmo ciclo após o seguinte halving. Seguindo esta estratégia, segundo a média dos três períodos entre halvings já ocorridos, haverá um retorno de 7934% sobre o investimento.

Este número se destaca especialmente se colocado contra o retorno médio de 2271% sobre o investimento no período entre os halvings. Portanto, a resposta à pergunta inicial sobre o problema é de que sim, há uma estratégia a ser definida que se baseie no tempo desde o último halving e o sentimento geral do mercado, e que se provou altamente eficiente ao longo do período de vida da criptomoeda.

Pasta do Drive com a base de dados e acesso ao
Google Colab:

[https://drive.google.com/drive/folders/1k3gZnvKpCexQXcyXgbC5qJqlXrXZBF4H?usp=s
haring](https://drive.google.com/drive/folders/1k3gZnvKpCexQXcyXgbC5qJqlXrXZBF4H?usp=s
haring)

