

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CAMPUS FLORIANÓPOLIS**

**Vinícius Felipe de Oliveira da Silva**

**Análise de Correlação do Bitcoin com NASDAQ-100 e Ouro**

**FLORIANÓPOLIS**

**2022**

VINÍCIUS FELIPE DE OLIVEIRA DA SILVA

Análise de Correlação do Bitcoin com NASDAQ-100 e Ouro

**Trabalho de Conclusão de Curso submetido à Universidade Federal de Santa Catarina, como requisito necessário para obtenção do grau de Bacharel em Engenharia Elétrica**

Florianópolis, Setembro de 2022

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Silva, Vinícius Felipe de Oliveira da  
Análise de correlação do Bitcoin com NASDAQ-100 e ouro /  
Vinícius Felipe de Oliveira da Silva ; orientador, Edson  
Cilos Vargas Júnior, 2022.  
61 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Centro Tecnológico,  
Graduação em Engenharia Elétrica, Florianópolis, 2022.

Inclui referências.

1. Engenharia Elétrica. 2. Bitcoin. 3. Ciência de  
Dados. 4. Estatística. I. Vargas Júnior, Edson Cilos. II.  
Universidade Federal de Santa Catarina. Graduação em  
Engenharia Elétrica. III. Título.

Vinícius Felipe de Oliveira da Silva

Título do TCC: Análise de Correlação do Bitcoin com NASDAQ-100 e Ouro

Este Trabalho Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharel em Engenharia Elétrica” e aceito, em sua forma final, pelo Curso de Graduação em Engenharia Elétrica.

Florianópolis, 06 de Setembro de 2022.



Documento assinado digitalmente  
**Miguel Moreto**  
Data: 16/09/2022 11:53:22-0300  
CPF: 948.850.100-63  
Verifique as assinaturas em <https://v.ufsc.br>

---

Prof. Miguel Moreto, Dr.  
Coordenador do Curso de Graduação em Engenharia Elétrica

**Banca Examinadora:**



Documento assinado digitalmente  
**EDSON CILOS VARGAS JUNIOR**  
Data: 07/09/2022 10:04:36-0300  
CPF: 088.267.299-13  
Verifique as assinaturas em <https://v.ufsc.br>

---

Prof. Edson Cilos Vargas Júnior, Dr.  
Universidade Federal de Santa Catarina



Documento assinado digitalmente  
**Eduardo Luiz Ortiz Batista**  
Data: 12/09/2022 09:05:52-0300  
CPF: 036.521.889-85  
Verifique as assinaturas em <https://v.ufsc.br>

---

Prof. Eduardo Luiz Ortiz Batista, Dr.  
Universidade Federal de Santa Catarina



Documento assinado digitalmente  
**Aldrovando Luis Azeredo Araujo**  
Data: 12/09/2022 17:51:05-0300  
CPF: 290.022.570-15  
Verifique as assinaturas em <https://v.ufsc.br>

---

Prof. Aldrovando Luis Azeredo Araújo, Dr.  
Universidade Federal de Santa Catarina

# Agradecimentos

Meus pais, por me apoiarem e me incentivarem incondicionalmente.

Meu irmão, por me motivar a ser alguém melhor a cada dia.

Kevin, Samuel, Vanderson, Vitor e Walter, por serem a minha família durante a universidade.

Luis, Kiyó, Rafael, Kalliandru e Miguel, por me apoiarem e me acompanharem nos meus *hobbies* e projetos pessoais.

Geórgia, Lara e Fusi, pelas risadas, momentos, vínculos e amizades.

Richard e Danilo, por me darem oportunidades e me ajudarem a me encontrar profissionalmente.



*“Você deve aproveitar os pequenos desvios.  
Ao máximo.  
Porque é onde você encontrará coisas mais importantes  
do que aquilo que você realmente deseja.”  
(Togashi Yoshihiro)*



# Resumo

Desde que o Bitcoin surgiu, ele tem sido muito utilizado como meio de investimento ou como meio de troca. Alguns dos seus defensores alegam que o ativo é uma reserva de valor, enquanto que os mais céticos o tratam como um ativo meramente especulativo. Para lançar luz ao tema, neste trabalho é feito o estudo da série histórica do Bitcoin e, através de técnicas de diferenciação de séries, é calculada a correlação do ativo com o NASDAQ-100 e com o ouro. Mediante algumas suposições estatísticas, com uma confiança de 95%, é demonstrado que em todos os três cenários avaliados (série completa do Bitcoin desde seu surgimento, desde 2017 e após o Covid-19) que o Bitcoin se assemelha mais ao NASDAQ-100, tendo uma correlação significativa ou até mesmo alta. Quando comparado ao ouro, o cenário mais provável é a correlação ser insignificativa, apesar de o período a partir de 2017 apresentar uma correlação significativa.

**Palavras-chave:** Bitcoin, NASDAQ-100, ouro, correlação, correlação de Pearson, correlação de Spearman, teste de Dickey-Fuller, intervalo de confiança, diferenciação fracionária.



# Abstract

Since Bitcoin emerged, it has been widely used as a medium of investment or as a medium of exchange. Some of its defenders claim that the asset is a store of value, while the most skeptical treat it as merely a speculative asset. To shed light on the topic, this work studies the Bitcoin historical series and, through series differentiation techniques, its correlation with NASDAQ-100 and gold. Considering some statistical assumptions, with a confidence of 95% it is shown that in all three scenarios evaluated (Bitcoin complete series since its emergence, since 2017 and after Covid-19) Bitcoin is more similar to the NASDAQ-100, with a significant or high correlation. When compared to gold, the most likely scenario is that the correlation is negligible, even though the period from 2017 onwards shows a significant correlation.

**Keywords:** Bitcoin, NASDAQ-100, gold, correlation, Pearson correlation, Spearman correlation, Dickey-Fuller test, confidence interval, fractional differentiation.



# Lista de ilustrações

Figura 1 – Preço do Bitcoin em função do dólar americano . . . . .	21
Figura 2 – Taxa de inflação da Europa e dos Estados Unidos . . . . .	21
Figura 3 – Anomalia da temperatura do ar . . . . .	22
Figura 4 – Exemplos de Correlação . . . . .	26
Figura 5 – O Quarteto de Anscombe . . . . .	27
Figura 6 – <i>Datassaurus Dozen</i> . . . . .	28
Figura 7 – Gráficos do consumo de queijo <i>per capita</i> e do número de pessoas que morreram por se entrelaçarem em seus lençóis nos Estados Unidos . . . . .	29
Figura 8 – Comparação de um modelo homocedástico com um heterocedástico . . . . .	31
Figura 9 – Valor das ações do NASDAQ-100 em dólares . . . . .	37
Figura 10 – Valor do retorno logarítmico do Bitcoin e das ações do Nasdaq-100 . . . . .	38
Figura 11 – Valor do ouro em dólares por onça . . . . .	39
Figura 12 – Valor do retorno logarítmico do Bitcoin e do ouro . . . . .	39
Figura 13 – Curvas da estatística-t (eixo esquerdo) e correlações de Pearson e Spearman (eixo direito) em comparação a diferenciação fracionária $d$ (eixo X) do Bitcoin . . . . .	40



# Lista de tabelas

Tabela 1 – Intervalos de correlação com Bitcoin no período completo . . . . .	41
Tabela 2 – Intervalos de correlação com Bitcoin a partir de 2017 . . . . .	41
Tabela 3 – Intervalos de correlação com Bitcoin a partir de 2020 . . . . .	42
Tabela 4 – Intervalos de correlação com Bitcoin no período completo . . . . .	42
Tabela 5 – Valores ótimos de $d$ para o período completo e intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária . . . . .	42
Tabela 6 – Intervalos de correlação com Bitcoin a partir de 2017 . . . . .	43
Tabela 7 – Valores ótimos de $d$ para o período a partir de 2017 e intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária . . . . .	43
Tabela 8 – Intervalos de correlação com Bitcoin a partir de 2020 . . . . .	44
Tabela 9 – Valores ótimos de $d$ para o período a partir de 2020 e intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária . . . . .	44



# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>19</b>
1.1	Bitcoin	20
1.2	Séries Temporais	21
1.3	Predição de Bitcoin	22
<b>2</b>	<b>METODOLOGIA</b>	<b>25</b>
2.1	Medidas Descritivas	25
2.2	Correlação de Pearson	25
2.3	Correlação de Spearman	27
2.4	Intervalo de Confiança	29
2.5	Séries Estacionárias	30
2.6	Heterocedasticidade	30
2.7	Série das Diferenças	30
2.8	Testes Unitários	31
2.9	Retorno Logarítmico	33
2.10	Diferença Fracionária	33
2.11	Implementação	34
<b>3</b>	<b>CORRELAÇÕES DO BITCOIN</b>	<b>37</b>
3.1	Bitcoin e NASDAQ-100	37
3.2	Bitcoin e Ouro	38
3.3	Diferença Fracionária nos ativos	38
3.4	Intervalo de Confiança	40
<b>4</b>	<b>RESULTADOS</b>	<b>41</b>
4.1	Correlações	41
4.1.1	Retorno logarítmico	41
4.1.1.1	Período Completo	41
4.1.1.2	Período a Partir de 2017	41
4.1.1.3	Período COVID-19	42
4.1.2	Diferença Fracionária	42
4.1.2.1	Período Completo	42
4.1.2.2	Período a Partir de 2017	43
4.1.2.3	Período COVID-19	43
<b>5</b>	<b>CONCLUSÃO</b>	<b>45</b>

<b>REFERÊNCIAS</b> .....	<b>47</b>
<b>APÊNDICES</b>	<b>51</b>
<b>APÊNDICE A – CÓDIGO DAS CORRELAÇÕES DE RETORNO LOGARITMO</b> .....	<b>53</b>
<b>APÊNDICE B – CÓDIGO DAS CORRELAÇÕES DE DIFERENÇA FRACIONÁRIA</b> .....	<b>57</b>

# 1 Introdução

A previsão de eventos é um fator importante para o planejamento e tomada de decisões. Utilizando modelos matemáticos, é possível realizar a inferência de ocorrência de um evento, através do cálculo de probabilidades [Frost 2019]. Modelos preditivos são cruciais para um projeto envolvendo séries temporais, como por exemplo:

- Na estimativa de demandas por insumos, o que pode afetar toda uma cadeia logística [Bueno 2016];
- Na previsão do tempo, que pode determinar a viabilidade da execução de uma obra ou até mesmo da demanda de produtos por parte de clientes [Brayshaw et al. 2020];
- Na criação de estratégias para compra e venda de ativos no mercado financeiro [Prado 2018].

A previsibilidade de um evento depende de diversos fatores, dentre eles [Hyndman e Athanasopoulos 2018]:

- O entendimento dos fatores que o afetam;
- Quantidade de dados disponíveis;
- Se a previsão afeta o que está sendo inferido.

Conforme [Hyndman e Athanasopoulos 2018], existem casos nas quais essas três condições são satisfeitas e, por isso, é possível construir modelos com alto poder preditivo. Um exemplo é a previsão do tempo. São conhecidos diversos fenômenos físicos, os quais influenciam no comportamento temporal. Há, também, uma quantidade enorme de dados disponíveis do histórico do tempo por muitos anos. Por último, o tempo não muda o seu comportamento conforme inferências são realizadas ao seu respeito.

Contudo, existem casos onde nem todos os critérios são satisfeitos, sendo o mercado financeiro um deles. Não existe um conhecimento exato e preciso sobre quais fatores afetam o mercado financeiro. Além disso, conforme previsões são feitas e ações são tomadas, o mercado pode ter um comportamento diferente do esperado, justamente pela interferência das previsões. O único item satisfeito é a quantidade de dados, os quais são abundantes <sup>1</sup>. Porém, Conforme [Hyndman e Athanasopoulos 2018], esse fator sozinho não é suficiente para garantir que o modelo adotado seja bom o suficiente.

---

<sup>1</sup> Apesar dessa informação ser verdadeira no geral, na prática é muito difícil coletar e armazenar todos os dados gerados no mercado, como por exemplo em livros de oferta de um ativo.

Dentre os vários ativos do mercado financeiro, estão as criptomoedas, sendo o Bitcoin o de maior destaque. Com a sua crescente valorização, esta criptomoeda tem sido considerada um investimento especulativo [Mehta, Agashe e Detroja 2021]. Desta forma, existe uma grande flutuação no valor do Bitcoin, implicando a necessidade de um bom modelo preditivo por partes dos investidores, sobretudo os que desejam maximizar o seu lucro através de compra e/ou venda do ativo.

Tendo em vista a relevância do ativo, assim como o crescente interesse pelo mesmo, neste trabalho será feita uma análise dos dados históricos do Bitcoin. Serão realizadas diversas análises estatísticas com a finalidade de fornecer uma descrição adequada dos dados. Além disso, serão testadas algumas hipóteses a fim de melhor compreender a série temporal do Bitcoin. Para atingir esse objetivo, será feita uma comparação do comportamento do Bitcoin com outros dois ativos: o índice NASDAQ-100 e o ouro. Por fim, mas não menos importante, este estudo poderá servir de base para trabalhos posteriores que busquem criar modelos preditivos, como por exemplo aqueles destinados a maximizar o lucro através de compra e/ou venda de Bitcoin.

## 1.1 Bitcoin

A ideia do Bitcoin foi descrita por Satoshi Nakamoto em seu *white paper* em 2008 [Nakamoto 2008]. Nele, Nakamoto descreveu uma nova "moeda" digital, *peer-to-peer*, anônima e totalmente descentralizada, ou seja, sem o controle de um indivíduo ou instituição [Antonopoulos 2017]. Até os dias atuais, ninguém sabe quem é, ou quem são, Satoshi Nakamoto.

Para que o Bitcoin possa funcionar sendo um ativo descentralizado, ele depende do consenso da comunidade, mais especificamente dos mineradores. Os mineradores recebem transações pendentes para serem confirmadas na rede e reúnem elas no chamado bloco. Para o bloco ser validado, é necessário resolver um problema computacional com o algoritmo SHA256 [Gilbert e Handschuh 2004]. Assim que um minerador resolver esse problema, ele transmite para o resto da rede o resultado com o bloco minerado. Se correto, o bloco é aceito como válido e um novo bloco é preparado para ser minerado. Como os mineradores precisam conferir o tempo todo as transações válidas e as pendentes, existe um consenso, o qual é a base para o Bitcoin ser descentralizado.

O Bitcoin começou a ser popularizado como meio de troca no tráfico de drogas [Mehta, Agashe e Detroja 2021] pela sua anonimidade. Com o tempo, o ativo foi se difundindo para demais usos, e o seu valor começou a crescer. De fato, o valor do Bitcoin cresceu tanto que o mesmo começou a ser visto mais como um investimento do que como uma moeda em si [Mehta, Agashe e Detroja 2021]. A Figura 1 mostra como o valor do Bitcoin subiu desde seu nascimento.

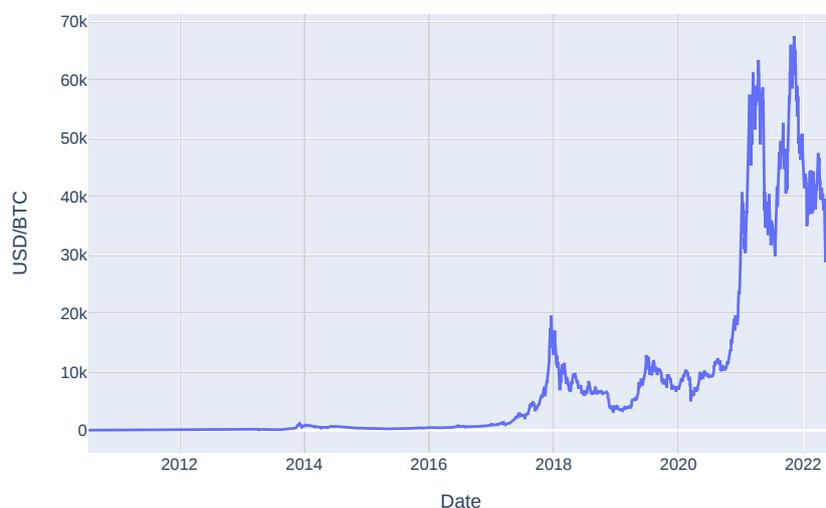


Figura 1 – Preço do Bitcoin em função do dólar americano

### *Acervo pessoal*

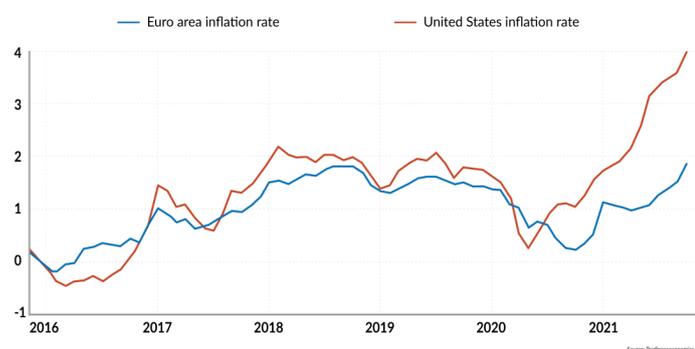


Figura 2 – Taxa de inflação da Europa e dos Estados Unidos

<https://tradingeconomics.com> acessado 09/06/22

## 1.2 Séries Temporais

Uma série temporal é qualquer conjunto de dados observados em função do tempo [Box, Jenkins e Reinsel 2008]. Diversas áreas utilizam séries temporais, desde economia, como inflação (Figura 2), até em ciências naturais, como temperatura (Figura 3).

Uma característica esperada das séries temporais é de que dados adjacentes observados possuem uma dependência entre si. Encontrar essa dependência é de extrema importância, pois permite compreender melhor o comportamento dos dados e, em alguns casos, fazer previsões para o futuro baseado nas observações presentes.

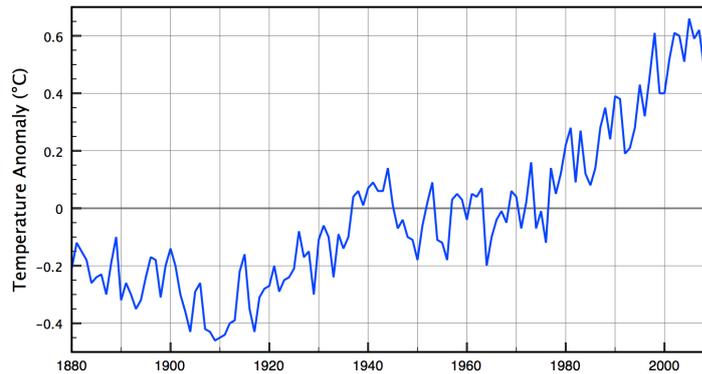


Figura 3 – Anomalia da temperatura do ar

<<https://serc.carleton.edu/details/images/50986.html>> acessado 09/06/22

### 1.3 Predição de Bitcoin

Para aqueles que tratam Bitcoin como investimento, tentar compreender melhor o seu comportamento é de extrema importância na tomada de decisão. Por exemplo, se o seu preço estiver subindo, é ideal que se segure a moeda na carteira (*HODL*) até o ponto em que ela chegue no pico, ponto no qual ela deve ser vendida. Em contraste, quando o valor do Bitcoin estiver descendo, o ideal é esperar até que ele chegue em seu pico negativo, para então comprá-lo.

Um dos exemplos mais básicos de tentativa de predição é a utilização de média móvel [Anupriya e Garg 2018]. Dentre as diversas variações de análises com média móvel, a mais simples é observar os momentos de transição do preço da moeda em relação a média móvel. Se o preço está em descida e atravessa a média móvel, está na hora de comprar. Quando o mesmo acontece com o preço em subida, vende-se a moeda.

Outra abordagem é feita tentando encontrar relações do preço do Bitcoin com demais ativos de valor no mercado. Um exemplo é o índice NASDAQ (*National Association of Securities Dealers Automated Quotations*), que é baseado em ações da bolsa de valores de Nova Iorque. Esse tipo de abordagem leva em consideração a ideia de que o Bitcoin se comporta como um ativo especulativo de risco [López-Cabarcos et al. 2021].

A NASDAQ é um mercado de ações pertencente a Nasdaq, Inc. de Nova Iorque baseado em diversas empresas. Ela é usada como referência mundial, sendo o NASDAQ-100, um índice baseado nas 100 maiores empresas não financeiras listadas em seu acervo, um dos indicadores mais usados no mercado (Figura 9).

Em contraste, outros autores comparam o Bitcoin a uma reserva de valor [Baur e Dimpfl 2021], a qual pode ser definida como ativos cujo valor é bem definido e não depreciado com o tempo [Mattke, Maier e Reis 2020], sendo o ouro um dos maiores exemplos [Dubey, Geanakoplos e Shubik 2003].

Entender a correlação do Bitcoin com outros ativos é importante para diminuir os riscos de investimento, investindo em outros ativos com comportamentos diferentes, diversificando o portfólio [Partovi, Caputo et al. 2004].

Existem resultados que mostram que o Bitcoin possui um comportamento cíclico em função do *Halving* [Meynkhard 2019]. O *Halving* é o fenômeno que ocorre quando a recompensa da mineração diminui pela metade depois de 210.000 blocos serem minerados, o que ocorre em um período aproximado de 4 anos.

Neste trabalho, será feita uma análise exploratória comparando o Bitcoin com demais ativos para entender se ele possui um comportamento mais parecido com um ativo de risco, como o NASDAQ-100, ou como uma reserva de valor, como o ouro. Para fazer tal análise, serão usadas ferramentas e conceitos estatísticos como correlação, séries estacionárias, heterocedasticidade, série das diferenças, teste unitário e retorno logaritmo.



## 2 Metodologia

Nesta seção, serão discutidos alguns dos conceitos e ferramentas utilizadas para a análise exploratória.

### 2.1 Medidas Descritivas

Esse trabalho assume que o leitor possua um conhecimento prévio de probabilidade e estatística, como os dois primeiros capítulos de [James 2004]. Neste trabalho é definido o valor esperado de uma variável pela equação

$$\mu = E[X] = \int_{\mathbb{R}} x dP, \quad (2.1)$$

sendo  $X$  uma variável aleatória e  $P$  a medida de probabilidade. A partir do valor esperado, é definida a variância por

$$\sigma_X^2 = V(X) = E[(X - \mu)^2] = \int_{\mathbb{R}} x^2 dP - \mu^2, \quad (2.2)$$

nas quais  $\sigma$  é o desvio padrão e  $V(X)$  a variância. A covariância  $\sigma_{XY}$  é definida pela equação

$$\text{cov}(X, Y) = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X \mu_Y. \quad (2.3)$$

Por fim, a correlação  $\rho_{XY}$  é definida por

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}. \quad (2.4)$$

### 2.2 Correlação de Pearson

A correlação de Pearson é uma variável que mede a relação linear entre dois sinais diferentes [Montgomery e Runger 2011]. O seu valor varia entre  $-1$  a  $1$  sendo esses valores considerados anti-correlação perfeita e correlação perfeita, respectivamente, além do  $0$ , o qual representa nenhuma correlação. Ela pode ser definida pelas seguintes expressões:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2.5)$$

sendo  $\rho$  a correlação de Pearson e  $\bar{x}$  a média amostral calculada por  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .

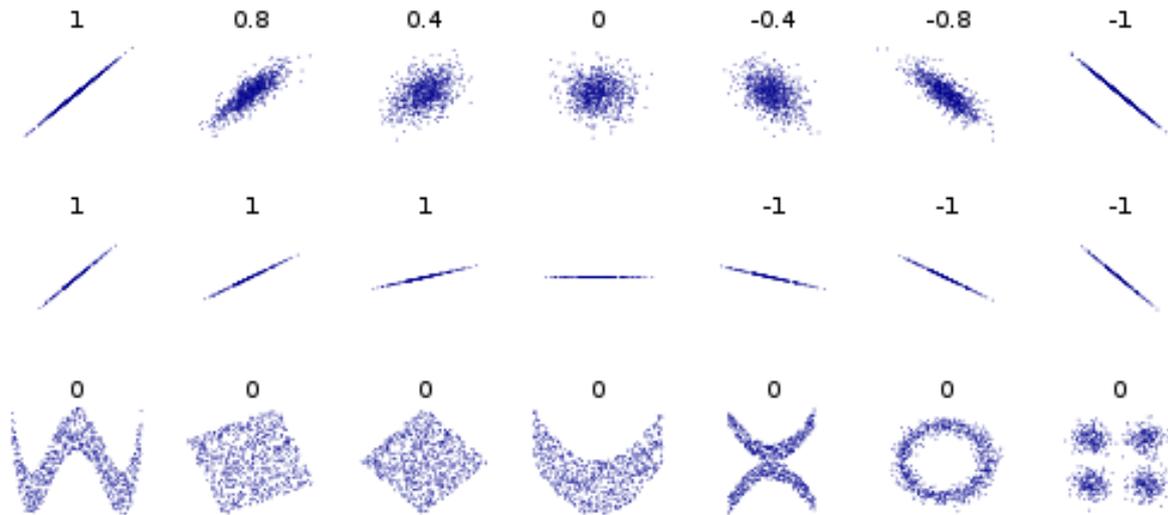


Figura 4 – Exemplos de Correlação

[https://en.wikipedia.org/wiki/File:Correlation\\_examples2.svg](https://en.wikipedia.org/wiki/File:Correlation_examples2.svg) acessado 30/04/22

Exemplos de correlação de Pearson podem ser vistos na Figura 4. A figura é interpretada como cada imagem representando uma distribuição de dados no plano XY de duas variáveis aleatórias, e o valor acima da distribuição representa o valor da correlação entre as duas variáveis. Nela, podemos ver a utilidade da correlação em encontrar relações e dependências lineares. Apesar disso, a correlação de Pearson possui algumas limitações. Assim como demonstrado na Figura 5, a correlação sozinha não dá informações suficientes sobre o sinal analisado. Conforme [Anscombe 1973], os quatro gráficos chamados de "Quarteto de Anscombe" possuem valores numéricos iguais para média e desvio padrão, tanto para o eixo  $X$  como  $Y$ , correlação de Pearson e coeficientes de curva de regressão linear. No entanto, as quatro curvas são completamente diferentes entre si. Isso mostra que, apesar da correlação ser uma ferramenta útil, ela deve ser usada em conjunto com demais técnicas para que seu resultado seja aproveitado da melhor forma. A Figura 6 mostra mais um exemplo dessa característica (mesma média, desvio padrão e correlação). Outra característica importante de se destacar é que correlação não implica em causalidade [Barrowman 2014]. Correlação indica apenas a forma que uma variável  $x$  se comporta em relação a uma outra variável  $y$ . Porém, essa relação não indica, necessariamente, que  $x$  depende de  $y$  ou vice-versa. É possível que sua relação seja apenas uma coincidência, também chamado de correlação espúria, ou que uma terceira variável afete ambas. A Figura 7 mostra um exemplo de correlação espúria.

Para evitar correlações espúrias e garantir um resultado válido e preciso, existem algumas restrições que a correlação de Pearson assume sobre os dados [Laerd Statistics 2020], sendo elas:

1. As variáveis  $X$  e  $Y$  são pareadas, isto é, denotando por  $(X, Y)$  o pareamento, as

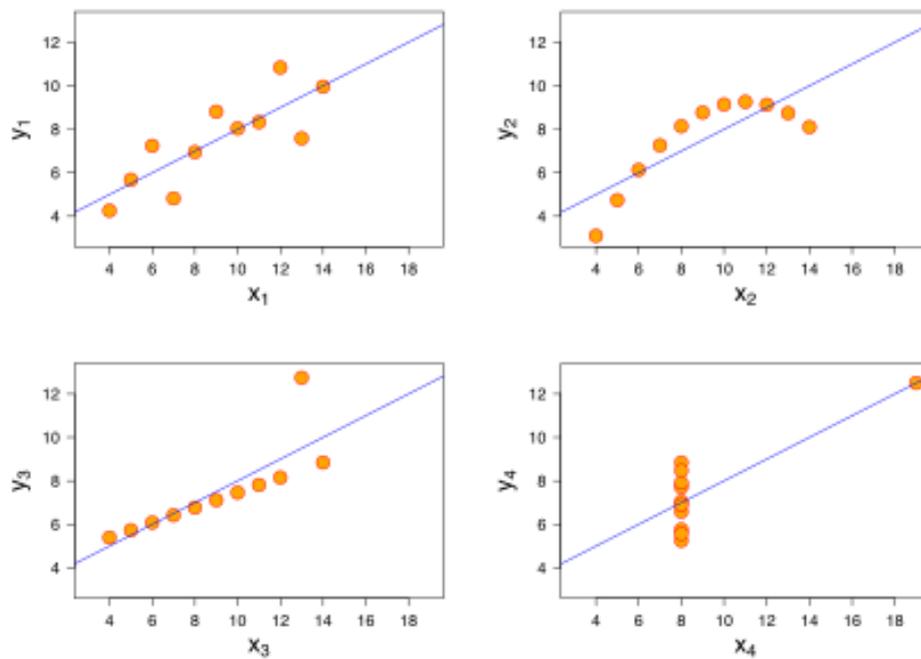


Figura 5 – O Quarteto de Anscombe

<[https://en.wikipedia.org/wiki/File:Anscombe%27s\\_quartet\\_3.svg](https://en.wikipedia.org/wiki/File:Anscombe%27s_quartet_3.svg)> acessado 30/04/22

instâncias de  $(XY)$  são elementos do conjunto  $D = (x_1, y_1), \dots, (x_n, y_n)$ , sendo que  $n$  é o tamanho da amostra;

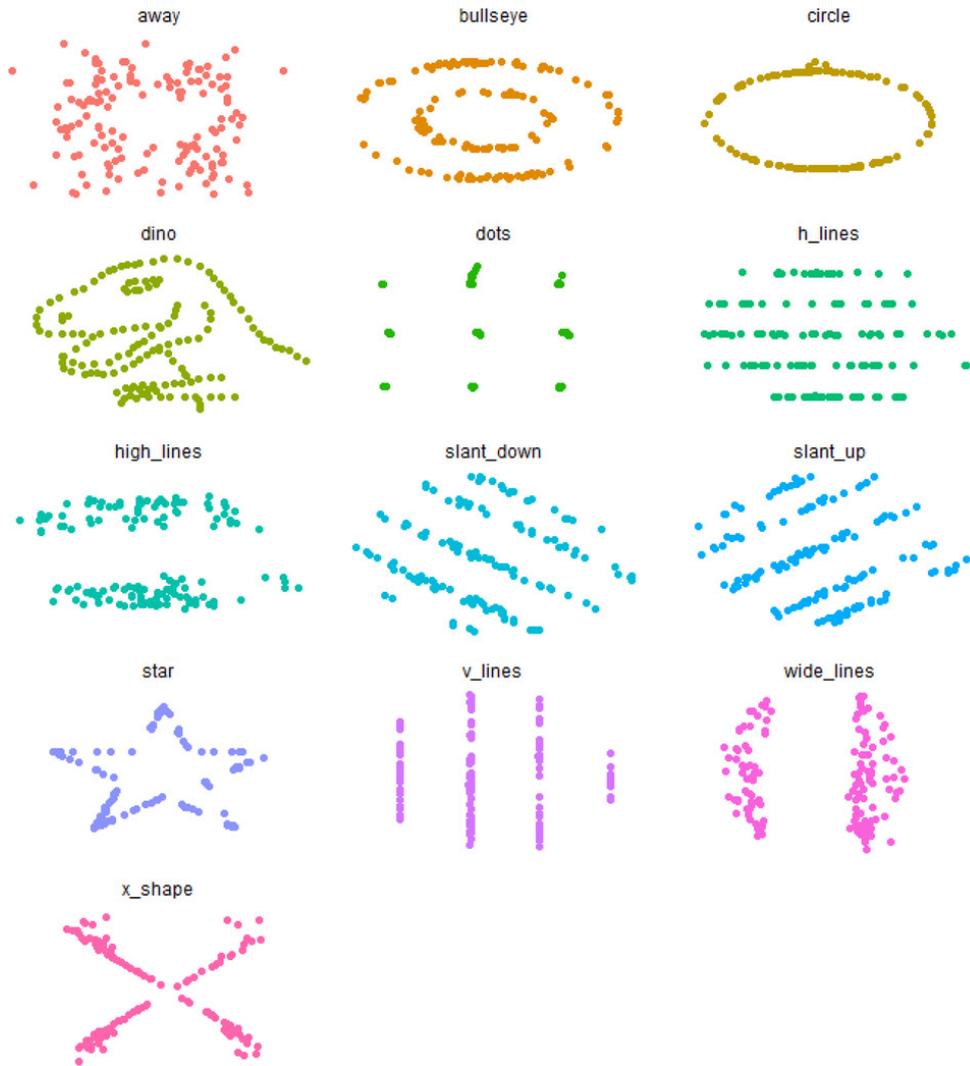
- Assumimos que  $D$  é identicamente e independentemente distribuído, sendo a sua distribuição a distribuição conjunta do par  $(X, Y)$ ;
- $(X, Y)$  segue uma distribuição normal bivariada.

Essas suposições são necessárias para fazer inferências a respeito da significância estatística e também do intervalo de confiança. Além disso, assumindo as condições 1, 2 e 3, através de uma formulação de máxima verossimilhança,  $r$  é a melhor estimativa para  $\rho$ .

## 2.3 Correlação de Spearman

A correlação de Spearman mede a força da relação monótona entre duas variáveis [Laerd Statistics 2020]. O seu valor varia entre -1 a 1 sendo esses valores considerados anti-correlação perfeita e correlação perfeita, respectivamente, além do 0, o qual representa nenhuma correlação.

Para calcular a correlação de Spearman, é criado um *ranking* das amostras das variáveis. O *ranking* é organizado de tal forma que, sendo  $x_1, \dots, x_n$  as instâncias de  $X$ , o maior  $x_k$  recebe o valor 1, o segundo maior recebe o ranking 2, até o menor  $x_k$ , o

Figura 6 – *Datassaurus Dozen*

[Murray e Wilson 2021]

qual recebe o valor  $n$ , sendo que  $k \in [1, n]$ . Assim, a correlação de Spearman é calculada encontrando o valor da correlação de Pearson entre os *rankings* das variáveis. No caso em que o *ranking* contém  $n$  valores inteiros distintos, a correlação de Spearman pode ser calculada pelas expressões:

$$\rho = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n d_i^2, \quad (2.6)$$

$$d_i = R(x_i) - R(y_i). \quad (2.7)$$

Sendo,  $\rho$  a correlação de Spearman e  $R(x_i)$  o *ranking* de  $x_i$ .

A correlação de Spearman é menos restritiva que a correlação de Pearson, na medida em que busca encontrar uma relação monótona, e não necessariamente linear como a de Pearson. As mesmas restrições de Pearson são aplicadas para Spearman, porém as restrições são para o *ranking* das variáveis aleatórias, não para elas em si.

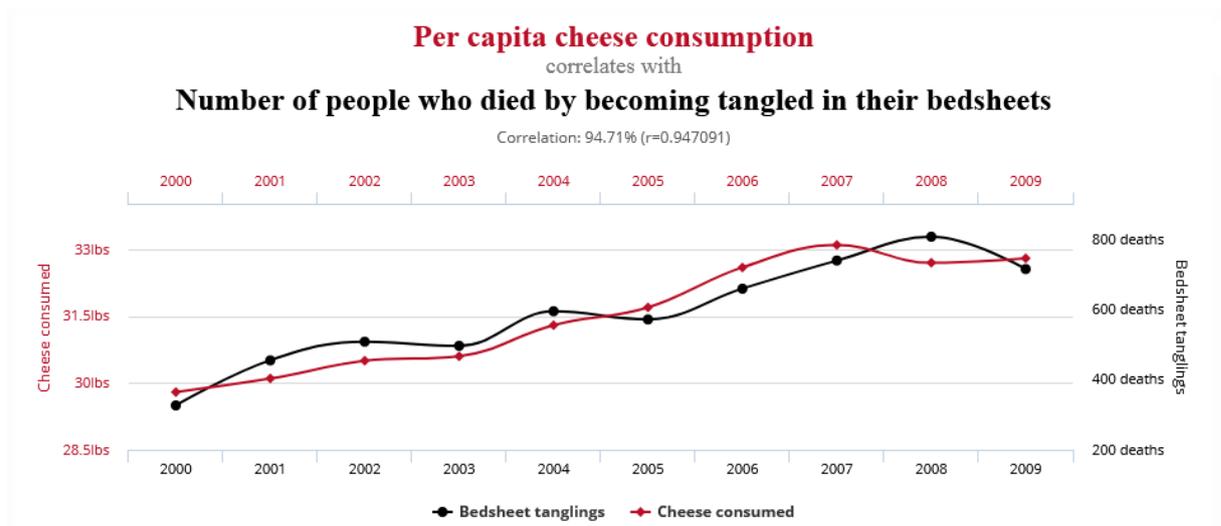


Figura 7 – Gráficos do consumo de queijo *per capita* e do número de pessoas que morreram por se entrelaçarem em seus lençóis nos Estados Unidos

<<https://www.tylervigen.com/spurious-correlations>> acessado 30/04/22

## 2.4 Intervalo de Confiança

As correlações de Pearson e Spearman são extrapolações feitas para poder calcular a correlação real (2.4), a qual é difícil de ser calculada na prática. Assim, existe um intervalo de confiança para garantir que o resultado encontrado atende dentro de um limite [Kutner 2005]. Assumindo as condições 1, 2, e 3 do final da seção 2.2, o intervalo de confiança pode ser encontrado a partir da transformada  $z$  de Fisher, definida por

$$z' = \frac{1}{2} \ln \left( \frac{1 + r_p}{1 - r_p} \right), \quad (2.8)$$

na qual  $r_p$  é a correlação de Pearson.

Quando o número de amostras  $n$  é suficientemente grande (como regra geral, assume-se  $n > 25$ ), a distribuição de  $z'$  se aproxima de uma distribuição normal e é possível aproximar a sua média e variância, respectivamente, por

$$E(z') = \zeta = \frac{1}{2} \ln \left( \frac{1 + \rho}{1 - \rho} \right), \quad (2.9)$$

$$\sigma^2(z') = \frac{1}{n - 3}, \quad (2.10)$$

sendo  $\rho$  a correlação.

Como a distribuição se aproxima de uma normal, o  $z$ -score pode ser calculado por

$$z = \frac{z' - \zeta}{\sigma(z')}. \quad (2.11)$$

Assim, para encontrar os limites inferior e superior para uma confiança de  $(1 - \alpha)\%$  para  $\zeta$ , são usadas as equações

$$\zeta_{inf} = z' - z \left( 1 - \frac{\alpha}{2} \right) \sigma(z'), \quad (2.12)$$

$$\zeta_{sup} = z' + z \left(1 - \frac{\alpha}{2}\right) \sigma(z'). \quad (2.13)$$

Dessa forma, tem-se os limites inferior e superior de *zeta*. Para encontrar os limites da correlação, basta substituir os resultados de (2.12) e (2.13) em (2.9).

O processo descrito representa o método para encontrar o intervalo de confiança para a correlação de Pearson. Para a correlação de Spearman  $r_s$ , a única diferença é o desvio padrão, aproximado por

$$\sigma^2(z') = \frac{1 + \frac{r_s^2}{2}}{n - 3}. \quad (2.14)$$

## 2.5 Séries Estacionárias

Seja  $\{X_t\}$  um processo estocástico e  $F_X(x_{t_1+\tau}, \dots, x_{t_n+\tau})$  a função de distribuição acumulada da distribuição de probabilidade conjunta incondicional de  $\{X_t\}$  para os períodos  $t_1 + \tau, \dots, t_n + \tau$ . Então,  $\{X_t\}$  é dito como estacionário se

$$F_X(x_{t_1+\tau}, \dots, x_{t_n+\tau}) = F_X(x_{t_1}, \dots, x_{t_n}), \forall \tau, t_1, \dots, t_n \in \mathbb{R} \wedge \forall n \in \mathbb{N}. \quad (2.15)$$

Como consequência, uma série estacionária é uma série cujas propriedades, como média e variância, não variam com o tempo [Hyndman e Athanasopoulos 2018]. Além disso, uma série estacionária não apresenta tendências, nem sazonalidade.

## 2.6 Heterocedasticidade

Dizemos que uma série  $\{X_t\}$  é homocedástica se a variância de  $\{X_t\}$  é constante para todo  $t$ , isto é, a variância da série é constante. A série será dita heterocedástica quando não é homocedástica [Greene 2003]. Por exemplo, na estimativa de máxima verossimilhança para o cálculo dos coeficientes da regressão linear, assume-se que a variável aleatória seja homocedástica. Além disso, uma série estacionária é também homocedástica.

## 2.7 Série das Diferenças

Existem maneiras de manipular séries não-estacionárias para que elas se tornem estacionárias. Uma forma é realizando a operação de diferenças. Seja  $\{X_t\}$  um processo estocástico com  $t$  entre 0 e  $T$ , a operação de diferenças pode ser definida por

$$x'_t = x_t - Bx_t = x_t - x_{t-1}, \quad (2.16)$$

$$B^k x_t = x_{t-k}, \forall t \in [k, T], \quad (2.17)$$

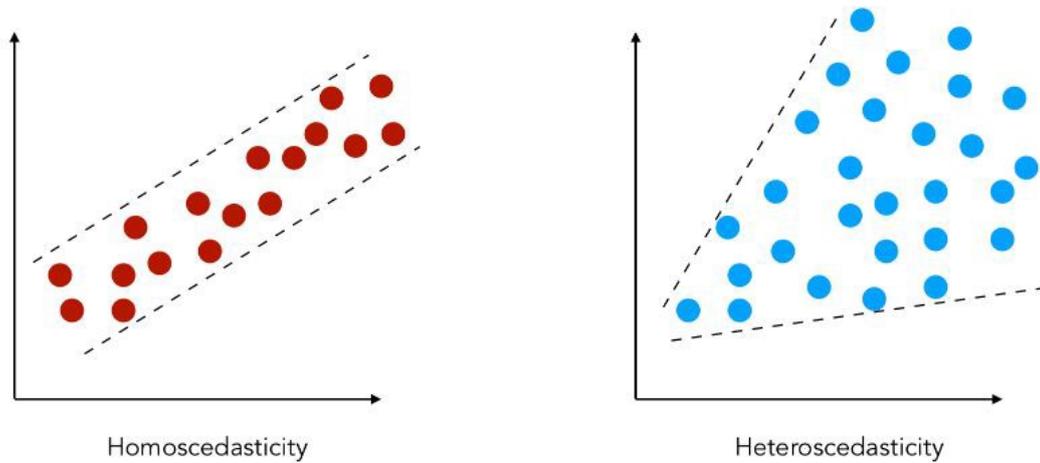


Figura 8 – Comparação de um modelo homocedástico com um heterocedástico

<<https://medium.com/riga-data-science-club/learn-heteroskedasticity-in-1-minute-42f678911389>>  
 acessado 02/06/22

sendo  $B$  chamado de operador *backshift* e  $T$  o número de amostras da série.

A equação (2.16) é uma série de diferença de primeira ordem. É possível fazer séries de diferenças de  $N$ -ordem caso a de primeira não resulte em uma série estacionária. Uma série de diferença de segunda ordem é definida por

$$x_t'' = x_t' - x_{t-1}' = x_t - 2x_{t-1} + x_{t-2}. \quad (2.18)$$

## 2.8 Testes Unitários

Testes unitários, também conhecidos como testes de raiz unitária, são testes utilizados para indicar a existência de uma raiz unitária em modelos autorregressivos. Tais testes são utilizados para ajudar a identificar se uma série é não-estacionária. Começando com o passeio aleatório, definido pela equação

$$y_t = By_t + \epsilon_t, \quad (2.19)$$

sendo  $B$  o operador *backshift* e  $\epsilon_t$  uma variável aleatória que representa o ruído. Manipulando (2.19), chegamos em

$$(1 - B)y_t = \epsilon_t, \quad (2.20)$$

caracterizada pela única raiz igual a 1, originando o nome de raiz unitária.

Dentre os diversos testes unitários, um dos mais simples e mais utilizados é o teste de Dickey-Fuller, o qual testa se  $\gamma = 1$  contra a hipótese alternativa  $\gamma < 1$ , considerando o modelo autorregressivo

$$y_t = \gamma y_{t-1} + \epsilon_t. \quad (2.21)$$

No caso em que  $\gamma = 1$  é uma solução de (2.21), tem-se uma raiz unitária e a variável é inferida como não-estacionária.

Compreendido o conceito de raiz unitária, será considerado o modelo mais geral

$$y_t - y_{t-1} = \mu\gamma + \beta\gamma t + (\gamma - 1)y_{t-1} + \epsilon_t, \quad (2.22)$$

sendo  $y$  a variável aleatória,  $\mu$  um parâmetro que indica uma direção (*drift*),  $\gamma$  um parâmetro artificial,  $\beta$  um coeficiente de tendência e  $\epsilon$  uma variável aleatória que representa o ruído.

Alguns casos especiais de (2.22) são, por exemplo:

- quando  $\mu = \beta = 0$  e  $\gamma = 1$ , representando um passeio aleatório;
- quando  $\beta = 0$  e  $\gamma = 1$ , representando um passeio aleatório com uma direção.

Procedimentos de inferência padrão, como regressão por mínimos quadrados, não são válidos para este cenário. Conforme  $\gamma$  se aproxima de 1, o valor estimado  $\hat{\gamma}$  através da regressão dos mínimos quadrado é enviesado para baixo, com variância amostral baixa [Evans e Savin 1981]. Isso implica que t-testes convencionais tem maior tendência a rejeitar incorretamente a hipótese  $\gamma = 1$ . Para testar a hipótese corretamente, deve-se usar uma lista de valores críticos gerados através de métodos de Monte Carlo [Greene 2003].

Na prática, tem-se um teste de hipótese nula  $H_0$ , o qual assume que a série possui uma raiz unitária e, portanto, é não-estacionária. A hipótese alternativa  $H_1$  sugere que a série não possui uma raiz unitária, sendo estacionária. Para rejeitar a hipótese, são usadas as equações

$$DF_\tau = \frac{\hat{\gamma} - \gamma_0}{SE(\hat{\gamma})}, \quad (2.23)$$

$$SE(\hat{\gamma}) = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}, \quad (2.24)$$

Sendo,  $DF_\tau$  a estatística-t,  $\hat{\gamma}$  o parâmetro estimado de  $\gamma$ ,  $\gamma_0$  o valor assumido de hipótese nula, e será igual a 1 nesse caso, dada a hipótese  $H_0$  e  $SE$  o erro padrão do parâmetro estimado  $\hat{\gamma}$  em relação a  $\gamma$ .

A estatística-t do teste é usada em conjunto com um valor crítico de confiança de grau  $c$ , com  $c$  entre 0 e 100. Se a estatística-t for menor que o valor de confiança de  $c$  %, a hipótese nula pode ser rejeitada, indicando que a série é estacionária.

Como o Dickey-Fuller é limitado, será utilizado o método de Dickey-Fuller aprimorado, definido por

$$y_t = \mu + \beta t + \gamma y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_p \Delta y_{t-p} + \epsilon_t,$$

$$\Delta y_t = y_t - y_{t-1}.$$

Para o caso ADF, as hipóteses  $H_0$  e  $H_1$ , a equação da estatística-t (2.23), e a lógica do teste são análogas ao do Dickey-Fuller [Greene 2003].

## 2.9 Retorno Logarítmico

O retorno logarítmico (ou log retorno) é uma operação realizada em séries temporais definida por

$$R_{log}(t) = \log(X_t/X_{t-1}) = \log(X_t) - \log(X_{t-1}). \quad (2.25)$$

Uma medida de correlação, como a de Pearson, aplicada a uma série temporal não-estacionária pode apresentar resultados espúrios, isto é, não ter significado prático [Yule 1926]. Uma das maneiras para contornar essa situação é aplicar o log retorno, que tipicamente torna a série estacionária e apresenta propriedades convenientes dependendo da distribuição dos dados.

## 2.10 Diferença Fracionária

Transformações, como a série das diferenças, tem o objetivo de gerar uma série estacionária a partir da original. Garantir a estacionariedade de uma série é importante quando se fala em previsões, pois muitos métodos de inferência partem do pressuposto que o sinal é estacionário [Prado 2018]. Porém, quando se aplica uma série das diferenças, é possível que o sinal resultante tenha perda de memória.

Para dados do mercado financeiro, tipicamente o sinal original possui memória, mas não é estacionário. O oposto pode ocorrer com log retorno ou diferenças inteiras da série. Esse é o dilema da estacionariedade vs. memória. Para transformar a série original em uma série estacionária mas, ao mesmo tempo, preservar o máximo de memória, pode-se utilizar a diferenciação fracionária [Prado 2018]. Partindo do teorema binomial para um número real  $d \geq 0$ , tem-se

$$(1 + x)^d = \sum_{k=0}^{\infty} \binom{d}{k} x^k, \quad (2.26)$$

assim, é possível obter a série da diferença de ordem  $d, \forall d \geq 0$ , a qual é definida por

$$Y_d = (1 - B)^d X, \quad (2.27)$$

$$(1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB + \frac{d(d-1)}{2!} B^2 - \frac{d(d-1)(d-2)}{3!} B^3 + \dots \quad (2.28)$$

sendo  $B$  o operador *backshift*.

A memória com um valor  $d$  real não-inteiro se mantém por

$$\tilde{X}_t = \sum_{k=0}^{\infty} \omega_k X_{t-k}, \quad (2.29)$$

$$\omega_0 = 1, \quad \omega_k = (-1)^k \prod_{i=0}^{k-1} \frac{d-i}{k!}. \quad (2.30)$$

Note que  $\prod_{i=0}^{k-1} \frac{d-i}{k!} = 0, \forall k > d$ , quando  $d$  é um número inteiro positivo. Portanto, para qualquer diferenciação inteira, a memória para  $k > d$  se perde.

É possível escrever os pesos  $\omega_k$  de forma recursiva considerando  $\omega_0 = 1$  como

$$\omega_k = \begin{cases} -\omega_{k-1} \frac{d-k+1}{k}, & \text{se } \omega_{k-1} \neq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (2.31)$$

A equação (2.31) mostra que os pesos convergem assintoticamente para zero, conforme  $k$  tende a infinito. Para  $k < d + 1$ , os pesos alternam entre valores positivos e negativos, enquanto que para  $k \geq d + 1$ , os pesos serão sempre negativos para  $d$  par, e positivos para  $d$  ímpar, arredondando  $d$  para baixo.

Para implementar computacionalmente, é necessário truncar pelo fato de existirem infinitos termos. Assim, são calculados pesos  $\omega_k$ , para  $k = 0, 1, \dots, T - l - 1$  com  $l$  sendo calculado a partir da equação

$$\lambda_l = \frac{\sum_{j=T-l}^T |\omega_j|}{\sum_{i=0}^{T-1} |\omega_i|}. \quad (2.32)$$

De (2.32), basta escolher o maior valor  $l$ , tal que  $\lambda_l \leq \tau$ , sendo  $\tau$  um valor de tolerância arbitrário entre 0 e 1.

## 2.11 Implementação

Os códigos do trabalho foram implementados em Python 3.7.13 e estão disponíveis no Apêndice. As bibliotecas usadas e suas respectivas versões são:

- Pandas 1.3.5 [The pandas development team 2020];
- Numpy 1.21.6 [Harris et al. 2020];
- Statsmodels 0.10.2 [Seabold e Perktold 2010];

- Plotly 5.5.0 [Plotly Technologies Inc. 2015];
- Matplotlib 3.2.2 [Hunter 2007].

Os dados históricos do Bitcoin foram retirados da [Coin Metrics](#), os dados da NASDAQ foram retirados da [Macrotrends](#) e os dados do ouro foram retirados da [Investing](#).



## 3 Correlações do Bitcoin

Nesta seção, serão discutidos os ativos a serem comparados com o Bitcoin, bem como a visualização de suas séries temporais. Os resultados finais serão apresentados no capítulo 4.

### 3.1 Bitcoin e NASDAQ-100

A NASDAQ (*National Association of Securities Dealers Automated Quotations*) é um mercado de ações pertencente a Nasdaq, Inc. de Nova Iorque baseado em diversas empresas. Ela é usada como referência mundial, sendo o NASDAQ-100 um índice baseado nas 100 maiores empresas não financeiras listadas em seu acervo, um dos indicadores mais usados no mercado (Figura 9).

Considerando que o Bitcoin possa ser considerado como um ativo especulativo [Mehta, Agashe e Detroja 2021], faz sentido compará-lo com o NASDAQ-100. A Figura 10 mostra o valor do retorno logarítmico do Bitcoin e o do NASDAQ-100.



Figura 9 – Valor das ações do NASDAQ-100 em dólares

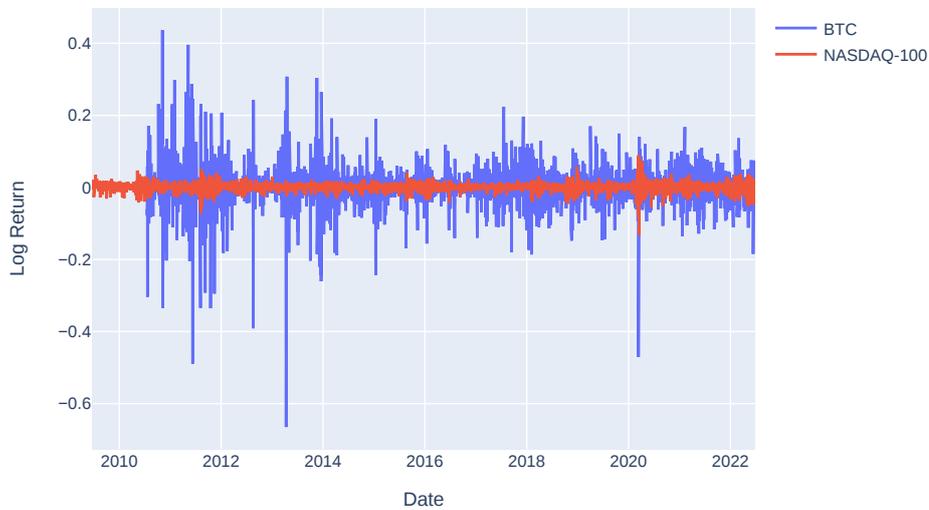


Figura 10 – Valor do retorno logarítmico do Bitcoin e das ações do Nasdaq-100

*Acervo pessoal*

## 3.2 Bitcoin e Ouro

O ouro tem sido usado como meio de troca por muito tempo [Spahn 2001]. Com o tempo, em função da dificuldade de transporte, seu uso foi sendo substituído pelas moedas fiduciárias. Atualmente, o ouro tem uma função maior como reserva de valor [Dubey, Geanakoplos e Shubik 2003].

A Figura 11 mostra o valor do ouro em USD/kg e a Figura 12 mostra o valor do retorno logarítmico do Bitcoin e do ouro.

## 3.3 Diferença Fracionária nos ativos

Baseando-se na seção 2.10, serão calculadas as séries de diferenças fracionárias com o melhor  $d$  para obter a estacionariedade e preservar o máximo de memória. Para encontrar o valor ótimo de  $d$ , será usado o teste ADF, descrito na sessão 2.8, com valor crítico de 5% como base, sendo o mesmo igual a  $-3,412$ . Aplicada a diferenciação fracionária, será calculada a estatística-t do sinal gerado. O menor valor de  $d$  cuja estatística-t seja menor que o valor crítico é o valor ótimo. Serão usado valores de  $d$  entre 0 e 1 com passo de 0,01 entre cada valor. O valor de  $\tau$  escolhido é de 0,05.

A figura 13 mostra a curva da estatística-t em função de  $d$ , sendo  $d$  a ordem da diferenciação fracionária. A linha horizontal representa o valor crítico. O primeiro valor de  $d$  depois que a curva da estatística-t atravessar o valor crítico é considerado o valor



Figura 11 – Valor do ouro em dólares por onça

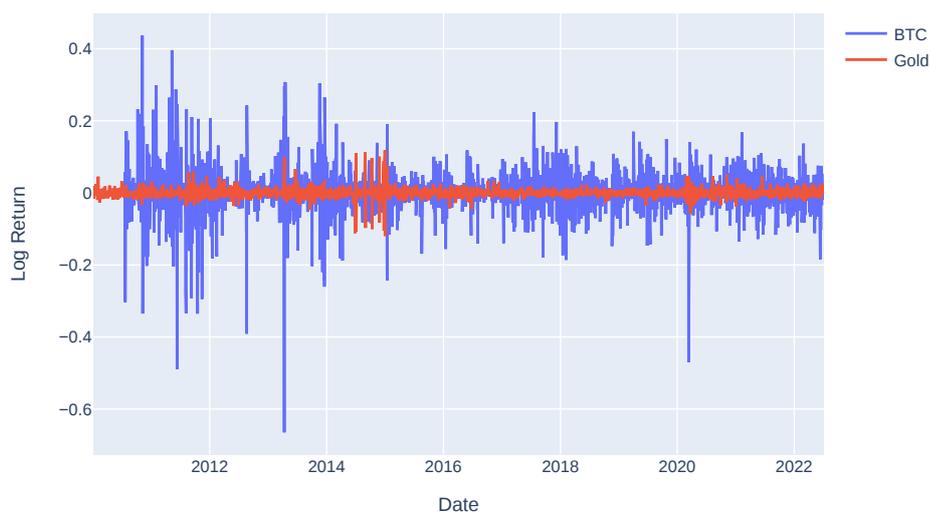
*Acervo pessoal*

Figura 12 – Valor do retorno logarítmico do Bitcoin e do ouro

*Acervo pessoal*

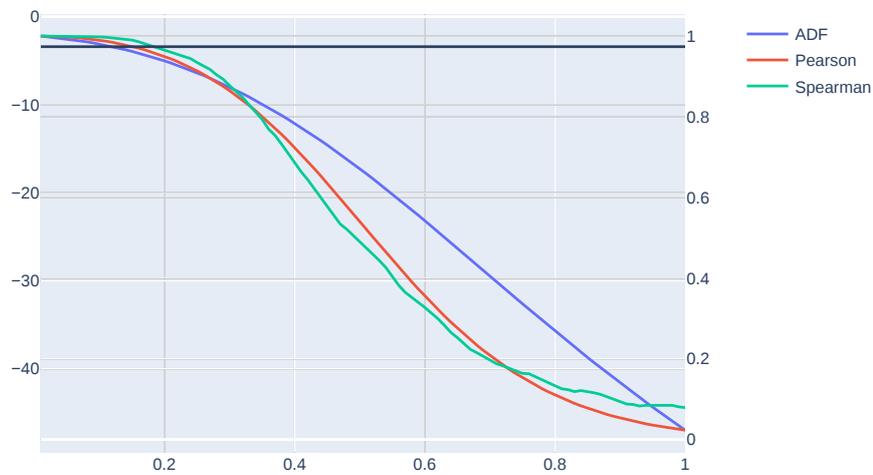


Figura 13 – Curvas da estatística-t (eixo esquerdo) e correlações de Pearson e Spearman (eixo direito) em comparação a diferenciação fracionária  $d$  (eixo X) do Bitcoin

#### *Acervo pessoal*

ótimo. Para efeito de comparação, foram inseridas as curvas de correlação de Pearson e de Spearman do sinal original com o sinal gerado a partir da diferenciação fracionária. Dessa forma, é possível ter um parâmetro para entender o quanto de memória do sinal original se manteve.

### 3.4 Intervalo de Confiança

O intervalo de confiança será calculado com limites de 95% de confiança, ou seja,  $\alpha$  será 0.05 nas equações 2.12 e 2.13.

## 4 Resultados

### 4.1 Correlações

Nesta seção, serão mostrados os resultados de correlação obtidos.

#### 4.1.1 Retorno logarítmico

##### 4.1.1.1 Período Completo

Na tabela 1, são mostrados os intervalos de confiança das correlações do retorno logarítmico do Bitcoin com o retorno logarítmico de cada índice respectivo. Tanto a NASDAQ-100 quanto o ouro possuem correlações baixas para esse período. Apesar do NASDAQ-100 apresentar uma correlação um pouco maior, ambas podem ser interpretadas como sendo irrelevantes.

Tabela 1 – Intervalos de correlação com Bitcoin no período completo

Indicador	Correlação de Pearson	Correlação de Spearman
NASDAQ-100	[0,091404; 0,161782]	[0,054202; 0,125293]
Ouro	[-0,0755404; -0,005462]	[-0,056712, 0,013454]

##### 4.1.1.2 Período a Partir de 2017

Na tabela 2, são mostrados os intervalos de confiança das correlações do retorno logarítmico do Bitcoin com o retorno logarítmico de cada índice respectivo a partir de 2017. Esse período foi escolhido por ser o período de primeiro grande pico no preço do Bitcoin e pelo fato do Bitcoin ter atingido certa maturidade, possivelmente menos volátil. A correlação com o NASDAQ-100 cresce consideravelmente para ambas as correlações nesse período, embora ainda seja baixa. Em contrapartida, a correlação com o ouro continua irrelevante.

Tabela 2 – Intervalos de correlação com Bitcoin a partir de 2017

Indicador	Correlação de Pearson	Correlação de Spearman
NASDAQ-100	[0,225455; 0,323207]	[0,143943; 0,246598]
Ouro	[-0,104508; -0,001878]	[-0,071785; 0,031105]

### 4.1.1.3 Período COVID-19

Na tabela 3, são mostrados os intervalos de confiança das correlações do retorno logarítmico do Bitcoin com o retorno logarítmico de cada índice respectivo a partir de 2020. Esse período foi escolhido por ser o período da pandemia da COVID-19, a qual impactou de maneira significativa na economia mundial. A correlação com o NASDAQ-100 teve mais um aumento considerável nesse período. O ouro continua tendo um valor insignificante.

Tabela 3 – Intervalos de correlação com Bitcoin a partir de 2020

Indicador	Correlação de Pearson	Correlação de Spearman
NASDAQ-100	[0,376248; 0,503133]	[0,288818; 0,430218]
Ouro	[-0,168154; -0,016601]	[-0,087923; 0,064927]

## 4.1.2 Diferença Fracionária

### 4.1.2.1 Período Completo

Na tabela 4, são mostrados os intervalos de confiança das correlações da diferença fracionária do Bitcoin com a diferença fracionária de cada índice respectivo. A tabela 5 mostra os valores ótimos de  $d$  encontrados para cada série, junto com os intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária de ordem  $d$ . Houve um aumento significativo na magnitude de ambas as correlações no período completo. A correlação com o NASDAQ-100 aproxima-se de uma correlação perfeita. A correlação com o ouro se tornou positiva e chegou em um valor significativo para a correlação de Pearson.

Tabela 4 – Intervalos de correlação com Bitcoin no período completo

Indicador	Correlação de Pearson	Correlação de Spearman
NASDAQ-100	[0,866970; 0,885096]	[0,925358; 0,937692]
Ouro	[0,212241; 0,291527]	[0,054634; 0,138704]

Tabela 5 – Valores ótimos de  $d$  para o período completo e intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária

	Bitcoin	NASDAQ-100	Ouro
$d$	0,12	0,15	0,06
Pearson	[0,983376; 0,985381]	[0,986107; 0,988020]	[0,985395; 0,987409]
Spearman	[0,995364; 0,996042]	[0,988990; 0,990814]	[0,979466; 0,982850]

#### 4.1.2.2 Período a Partir de 2017

Na tabela 6, são mostrados os intervalos de confiança das correlações da diferença fracionária do Bitcoin com a diferença fracionária de cada índice respectivo a partir de 2017. A tabela 7 mostra os valores ótimos de  $d$  encontrados para cada série, junto com os intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária de ordem  $d$ . Em contraste com o retorno logarítmico, a correlação com o NASDAQ-100 diminuiu durante esse período, apesar de se manter alta. A correlação com o ouro subiu significativamente, ficando próxima da correlação com o NASDAQ-100.

Tabela 6 – Intervalos de correlação com Bitcoin a partir de 2017

Indicador	Correlação de Pearson	Correlação de Spearman
NASDAQ-100	[0,755280; 0,800439]	[0,727473; 0,783133]
Ouro	[0,527814; 0,612094]	[0,578603; 0,662409]

Tabela 7 – Valores ótimos de  $d$  para o período a partir de 2017 e intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária

	Bitcoin	NASDAQ-100	Ouro
$d$	0,19	0,23	0,17
Pearson	[0,928854; 0,940796]	[0,857260; 0,884687]	[0,980096; 0,984102]
Spearman	[0,938206; 0,950511]	[0,796648; 0,839992]	[0,973075; 0,979492]

#### 4.1.2.3 Período COVID-19

Na tabela 8, são mostrados os intervalos de confiança das correlações da diferença fracionária do Bitcoin com a diferença fracionária de cada índice respectivo a partir de 2020. A tabela 9 mostra os valores ótimos de  $d$  encontrados para cada série, junto com os intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária de ordem  $d$ . A correlação com o NASDAQ-100 é menor nesse período em relação aos demais, apesar de se manter relativamente alta. O resultado com o ouro sofreu uma grande alteração, diminuindo novamente seu valor e o intervalo de confiança calculado não é suficiente para garantir que a correlação do ouro com o Bitcoin seja significativa.

Tabela 8 – Intervalos de correlação com Bitcoin a partir de 2020

Indicador	Correlação de Pearson	Correlação de Spearman
NASDAQ-100	[0,424206; 0,555559]	[0,422835; 0,561524]
Ouro	[0,042569; 0,230128]	[0,038801; 0,227369]

Tabela 9 – Valores ótimos de  $d$  para o período a partir de 2020 e intervalos de correlações de Pearson e Spearman de cada série em relação a sua respectiva série de diferença fracionária

	Bitcoin	NASDAQ-100	Ouro
$d$	0,27	0,25	0,13
Pearson	[0,770563; 0,740618]	[0,290428; 0,440306]	[0,873390; 0,909012]
Spearman	[0,796719; 0,848649]	[0,378702; 0,523280]	[0,807987; 0,867854]

## 5 Conclusão

Foi feito um estudo sobre o Bitcoin para identificar se o mesmo se assemelha mais a uma reserva de valor, como o ouro, ou como um ativo especulativo, como o NASDAQ-100. As correlações foram significativamente maiores em todos os períodos com o NASDAQ-100, indicando uma semelhança maior com um ativo especulativo, principalmente para períodos mais recentes, em especial o período pós COVID-19. As correlações com o ouro têm magnitude baixa, indicando que os dois ativos não possuem qualquer dependência entre si. Não havia expectativa sobre os resultados por parte do autor.

Além disso, houve uma diferença considerável para as correlações com a transformação das séries pelas diferença fracionária, obtendo resultados mais expressivos. Isso indica que a diferença fracionária é uma técnica que consegue preservar memória ao mesmo tempo em que transforma a série em estacionária de maneira efetiva, e, provavelmente, deve-se obter melhores resultados em demais séries além das apresentadas no trabalho.

Este trabalho pode servir como base para estudos futuros sobre a natureza do Bitcoin e seu comportamento. Com uma análise exploratória maior utilizando demais parâmetros e ferramentas seria possível encontrar mais relações e, possivelmente, ajudar na predição do preço futuro do ativo.



# Referências

- ANSCOMBE, F. J. Graphs in statistical analysis. *The American Statistician*, v. 27, n. 1, p. 17–21, 1973. Citado na página 26.
- ANTONOPOULOS, A. M. *Mastering Bitcoin*. 2. ed. [S.l.]: O'Reilly Media, inc, 2017. Citado na página 20.
- ANUPRIYA; GARG, S. Autoregressive integrated moving average model based prediction of bitcoin close price. In: *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*. [S.l.: s.n.], 2018. p. 473–478. Citado na página 22.
- BARROWMAN, N. Correlation, causation, and confusion. *The New Atlantis*, n. 43, p. 23–44, 2014. Citado na página 26.
- BAUR, D. G.; DIMPFL, T. The volatility of bitcoin and its role as a medium of exchange and a store of value. *Empirical Economics*, Springer, v. 61, n. 5, p. 2663–2683, 2021. Citado na página 22.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. *Time Series Analysis - Forecasting and Control*. 4. ed. [S.l.]: John Wiley & Sons, Inc, 2008. Citado na página 21.
- BRAYSHAW, D. J. et al. Quantifying the potential for improved management of weather risk using sub-seasonal forecasting: The case of uk telecommunications infrastructure. *Meteorological Applications*, v. 27, n. 1, p. e1849, 2020. Disponível em: <<https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.1849>>. Citado na página 19.
- BUENO, G. C. *Previsão de demanda de insumos relacionados a produção de café em máquinas automáticas*. 2016. Disponível em: <<https://repositorio.ufsm.br/handle/1/2694>>. Citado na página 19.
- DUBEY, P.; GEANAKOPOLOS, J.; SHUBIK, M. Is gold an efficient store of value? *Economic Theory*, Springer, v. 21, n. 4, p. 767–782, 2003. Citado 2 vezes nas páginas 22 e 38.
- EVANS, G.; SAVIN, N. E. Testing for unit roots: 1. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 753–779, 1981. Citado na página 32.
- FROST, J. *Regression Analysis*. 1. ed. [S.l.]: Statistics By Jim Publishing, 2019. Citado na página 19.
- GILBERT, H.; HANDSCHUH, H. Security analysis of sha-256 and sisters. In: MATSUI, M.; ZUCCHERATO, R. J. (Ed.). *Selected Areas in Cryptography*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 175–193. ISBN 978-3-540-24654-1. Citado na página 20.
- GREENE, W. *Econometric analysis fifth edition*. 01 2003. Citado 3 vezes nas páginas 30, 32 e 33.

- HARRIS, C. R. et al. Array programming with NumPy. *Nature*, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020. Disponível em: <<https://doi.org/10.1038/s41586-020-2649-2>>. Citado na página 34.
- HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007. Citado na página 35.
- HYNDMAN, R.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. 2. ed. OTexts: Melbourne, Australia, 2018. Disponível em: <<https://otexts.com/fpp2/index.html>>. Citado 2 vezes nas páginas 19 e 30.
- JAMES, B. *Probabilidade: um curso em nível intermediário*. Instituto de Matemática Pura e Aplicada, 2004. (Projeto Euclides). ISBN 9788524401015. Disponível em: <<https://books.google.com.br/books?id=xT71HAAACAAJ>>. Citado na página 25.
- KUTNER, M. *Applied Linear Statistical Models*. McGraw-Hill Irwin, 2005. (McGrwa-Hill international edition). ISBN 9780071122214. Disponível em: <<https://books.google.com.br/books?id=0xqCAAACAAJ>>. Citado na página 29.
- LAERD STATISTICS. Pearson’s product moment correlation. *Statistical tutorials and software guides*, 2020. Disponível em: <<https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>>. Citado na página 26.
- LAERD STATISTICS. Spearman’s rank-order correlation. *Statistical tutorials and software guides*, 2020. Disponível em: <<https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-guide.php>>. Citado na página 27.
- LÓPEZ-CABARCOS, M. Á. et al. Bitcoin volatility, stock market and investor sentiment. are they connected? *Finance Research Letters*, Elsevier, v. 38, p. 101399, 2021. Citado na página 22.
- MATTKE, J.; MAIER, C.; REIS, L. Is cryptocurrency money? three empirical studies analyzing medium of exchange, store of value and unit of account. In: *Proceedings of the 2020 on Computers and People Research Conference*. [S.l.: s.n.], 2020. p. 26–35. Citado na página 22.
- MEHTA, N.; AGASHE, A.; DETROJA, P. *Blockchain Bubble or Revolution: The Future of Bitcoin, Blockchains, and Cryptocurrencies*. 2. ed. Paravane Ventures, 2021. Disponível em: <<https://bubbleorrevolution.com/>>. Citado 2 vezes nas páginas 20 e 37.
- MEYNKHARD, A. Fair market value of bitcoin: Halving effect. *Investment Management and Financial Innovations*, v. 16, n. 4, p. 72–85, 2019. Citado na página 23.
- MONTGOMERY, D. C.; RUNGER, G. C. *Applied Statistics and Probability for Engineers*. 5. ed. [S.l.]: John Wiley & Sons, Inc, 2011. Citado na página 25.
- MURRAY, L. L.; WILSON, J. G. Generating data sets for teaching the importance of regression analysis. *Decision Sciences Journal of Innovative Education*, v. 19, n. 2, p. 157–166, 2021. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/dsji.12233>>. Citado na página 28.
- NAKAMOTO, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*. 2008. Disponível em: <<https://bitcoin.org/bitcoin.pdf>>. Citado na página 20.

- PARTOVI, M. H.; CAPUTO, M. et al. Principal portfolios: Recasting the efficient frontier. *Economics Bulletin*, AccessEcon, v. 7, n. 3, p. 1–10, 2004. Citado na página 23.
- PLOTLY TECHNOLOGIES INC. *Collaborative data science*. Montreal, QC: Plotly Technologies Inc., 2015. Disponível em: <<https://plot.ly>>. Citado na página 35.
- PRADO, M. Lopez de. *Advances in Financial Machine Learning*. 1. ed. [S.l.]: John Wiley & Sons, Inc, 2018. Citado 2 vezes nas páginas 19 e 33.
- SEABOLD, S.; PERKTOLD, J. statsmodels: Econometric and statistical modeling with python. In: *9th Python in Science Conference*. [S.l.: s.n.], 2010. Citado na página 34.
- SPAHN, H.-P. *From gold to euro: On monetary theory and the history of currency systems*. [S.l.]: Springer Science & Business Media, 2001. Citado na página 38.
- THE PANDAS DEVELOPMENT TEAM. *pandas-dev/pandas: Pandas*. Zenodo, 2020. Disponível em: <<https://doi.org/10.5281/zenodo.3509134>>. Citado na página 34.
- YULE, G. U. Why do we sometimes get nonsense-correlations between time-series?—a study in sampling and the nature of time-series. *Journal of the royal statistical society*, JSTOR, v. 89, n. 1, p. 1–63, 1926. Citado na página 33.



# Apêndices



# APÊNDICE A – Código das correlações de retorno logaritmo

```
# https://github.com/Vinicius-Felipe/tcc-bitcoin

import pandas as pd
import numpy as np
import math
from scipy.stats import norm

def log_return(series):
    series_log = np.log(series).diff()
    series_log = series_log.dropna()
    return series_log

def period(series, year):
    series_period = series[series.index.year >= year]
    return series_period

def interval_confidence(series_a, series_b, confidence = 0.95, method = 'pearson'):
    confidence = 0.95
    r = series_a.corrwith(series_b, method = method)[0]
    num = len(series_a.index.intersection(series_b.index))
    if method == 'pearson':
        stderr = 1.0 / math.sqrt(num - 3)
    elif method == 'spearman':
        stderr = math.sqrt((1 + r*r/2)/(num - 3))
    delta = norm.ppf((1 + confidence)/2) * stderr
    lower = math.tanh(math.atanh(r) - delta)
    upper = math.tanh(math.atanh(r) + delta)
    return lower, upper

btc = pd.read_csv('drive/My Drive/TCC Vinicius/btc.csv')
btc = btc.rename(columns={"time": "date"})
btc = btc.set_index('date')
btc_price = pd.DataFrame(btc['PriceUSD'], index = btc.index)
btc_price = btc_price.dropna()
btc_price.index = pd.to_datetime(btc_price.index)
btc_price_log = log_return(btc_price)
btc_price_log_2017 = period(btc_price_log, 2017)
btc_price_log_2020 = period(btc_price_log, 2020)
```

```

nasdaq = pd.read_csv('drive/My Drive/TCC Vinicius/nasdaq-100.csv',
    ↪ on_bad_lines='skip')
nasdaq = nasdaq.set_index('date')
nasdaq.index = pd.to_datetime(nasdaq.index)
nasdaq_log = log_return(nasdaq)
nasdaq_log = nasdaq_log[nasdaq_log.columns[0]]

gold = pd.read_csv('drive/My Drive/TCC Vinicius/gold.csv')
gold = gold.rename(columns={"Date": "date"})
gold = gold.set_index('date')
gold.index = pd.to_datetime(gold.index)
gold['Price'] = gold['Price'].str.replace(',', '').astype(float)
gold_log = log_return(gold['Price'])

def print_correlation():

    print('Correlacao de Pearson BTC x NASDAQ-100 período inteiro:',
    ↪ btc_price_log.corrwith(nasdaq_log)[0], interval_confidence(btc_price_log,
    ↪ nasdaq_log))
    print('Correlacao de Spearman BTC x NASDAQ-100 período inteiro:',
    ↪ btc_price_log.corrwith(nasdaq_log, method = 'spearman')[0],
    ↪ interval_confidence(btc_price_log, nasdaq_log, method = 'spearman'))
    print('Correlacao de Pearson BTC x ouro período inteiro:',
    ↪ btc_price_log.corrwith(gold_log)[0], interval_confidence(btc_price_log,
    ↪ gold_log))
    print('Correlacao de Spearman BTC x ouro período inteiro:',
    ↪ btc_price_log.corrwith(gold_log, method = 'spearman')[0],
    ↪ interval_confidence(btc_price_log, gold_log, method = 'spearman'))

    print('')

    print('Correlacao de Pearson BTC x NASDAQ-100 a partir de 2017:',
    ↪ btc_price_log_2017.corrwith(nasdaq_log)[0],
    ↪ interval_confidence(btc_price_log_2017, nasdaq_log))
    print('Correlacao de Spearman BTC x NASDAQ-100 a partir de 2017:',
    ↪ btc_price_log_2017.corrwith(nasdaq_log, method = 'spearman')[0],
    ↪ interval_confidence(btc_price_log_2017, nasdaq_log, method = 'spearman'))
    print('Correlacao de Pearson BTC x ouro a partir de 2017:',
    ↪ btc_price_log_2017.corrwith(gold_log)[0],
    ↪ interval_confidence(btc_price_log_2017, gold_log))
    print('Correlacao de Spearman BTC x ouro a partir de 2017:',
    ↪ btc_price_log_2017.corrwith(gold_log, method = 'spearman')[0],
    ↪ interval_confidence(btc_price_log_2017, gold_log, method = 'spearman'))

    print('')

```

```
print('Correlacao de Pearson BTC x NASDAQ-100 a partir de 2020:',
      ↪ btc_price_log_2020.corrwith(nasdaq_log)[0],
      ↪ interval_confidence(btc_price_log_2020, nasdaq_log))
print('Correlacao de Spearman BTC x NASDAQ-100 a partir de
      ↪ 2020:', btc_price_log_2020.corrwith(nasdaq_log, method = 'spearman')[0],
      ↪ interval_confidence(btc_price_log_2020, nasdaq_log, method = 'spearman'))
print('Correlacao de Pearson BTC x ouro a partir de 2020:',
      ↪ btc_price_log_2020.corrwith(gold_log)[0],
      ↪ interval_confidence(btc_price_log_2020, gold_log))
print('Correlacao de Spearman BTC x ouro a partir de 2020:',
      ↪ btc_price_log_2020.corrwith(gold_log, method = 'spearman')[0],
      ↪ interval_confidence(btc_price_log_2020, gold_log, method = 'spearman'))
return

if __name__ == '__main__':
    print_correlation()
```



# APÊNDICE B – Código das correlações de diferença fracionária

```
# https://github.com/Vinicius-Felipe/tcc-bitcoin

import pandas as pd
import numpy as np
from statsmodels.tsa.stattools import adfuller
import plotly.graph_objects as go
pd.options.plotting.backend = "plotly"
import math
from scipy.stats import norm

def period(series, year):
    series_period = series[series.index.year >= year]
    return series_period

def getWeights(d,size):
    # thres > 0 drops insignificant weights
    w=[1.]
    for k in range(1,size):
        w_ = -w[-1]/k*(d-k+1)
        w.append(w_)
    w=np.array(w[:-1]).reshape(-1,1)
    return w

def fracDiff (series, d, thres=.01):
    """
    Increasing width window, with treatment of Nas
    Note 1: For thres=1, nothing is skipped.
    Note 2: d can be any positive fractional, not necessarily bounded [0,1].
    """
    #1) Compute weights for the longest series
    w=getWeights (d, series.shape[0])
    #2) Determine initial calcs to be skipped based on weight-loss threshold
    w_ = np.cumsum(abs(w))
    w_/=w_[-1]
    skip = w_[w_>thres].shape[0]
    #3) Apply weights to values
    df = {}
    for name in series.columns:
        seriesF, df_ = series[[name]].fillna(method='ffill').dropna(),
        ↪ pd.Series(dtype='float64')
```

```

for iloc in range(skip, seriesF.shape[0]):
    loc=seriesF.index[iloc]
    if not np.isfinite(series.loc[loc, name]) :continue #exclude NAs
    df_[loc] = np.dot(w[-(iloc + 1):,:].T, seriesF.loc[:loc])[0,0]
    df[name] =df_.copy(deep=True)
df = pd.concat(df,axis=1)
return df

def adf_curve(ser):
    range_idx = int(100)
    pearson = []
    adf = []
    spearman = []
    d_opt = None
    for d in range(range_idx):
        frac = fracDiff(ser, (d+1)/range_idx, 0.05)
        pearson.append(frac.corrwith(ser)[0])
        adf_t = adfuller(frac.iloc[:,0], maxlag = 1, regression = 'ct', autolag = None)[0]
        adf.append(adf_t)
        spearman.append(frac.corrwith(ser, method = 'spearman')[0])
        if d_opt is None:
            if adf_t < adf_95:
                d_opt = (d + 1)/100
                pearson_opt = [frac.corrwith(ser)[0], interval_confidence(frac, ser)]
                spearman_opt = [frac.corrwith(ser, method = 'spearman')[0],
                    ↪ interval_confidence(frac, ser, method = 'spearman')]
    from plotly.subplots import make_subplots
    fig = make_subplots(specs=[[{"secondary_y": True}]])
    fig.add_trace(go.Scatter(x=np.linspace(1,range_idx,range_idx)/range_idx , y=adf ,
        ↪ name="ADF",), secondary_y=False,)
    fig.add_trace(go.Scatter(x=np.linspace(1,range_idx,range_idx)/range_idx , y=pearson ,
        ↪ name="Pearson",), secondary_y=True,)
    fig.add_trace(go.Scatter(x=np.linspace(1,range_idx,range_idx)/range_idx , y=spearman
        ↪ , name="Spearman",), secondary_y=True,)
    fig.add_hline(y=adf_95)
    fig.update_yaxes(title_text="ADF", secondary_y=False)
    fig.update_yaxes(title_text="Correlation", secondary_y=True)
    fig.update_xaxes(title_text="d")
    fig.show()
    return d_opt, pearson_opt, spearman_opt

def interval_confidence(series_a, series_b, confidence = 0.95, method = 'pearson'):
    confidence = 0.95
    r = series_a.corrwith(series_b, method = method)[0]
    num = len(series_a.index.intersection(series_b.index))
    if method == 'pearson':
        stderr = 1.0 / math.sqrt(num - 3)

```

```
elif method == 'spearman':
    stderr = math.sqrt( (1 + r*r/2)/(num - 3))
    delta = norm.ppf( (1+ confidence)/2) * stderr
    lower = math.tanh(math.atanh(r) - delta)
    upper = math.tanh(math.atanh(r) + delta)
    return lower, upper

btc = pd.read_csv('drive/My Drive/TCC Vinicius/btc.csv')
btc = btc.rename(columns={"time": "date"})
btc = btc.set_index('date')
btc_price = pd.DataFrame(btc['PriceUSD'], index = btc.index)
btc_price = btc_price.dropna()
btc_price.index = pd.to_datetime(btc_price.index)
btc_price_2017 = period(btc_price, 2017)
btc_price_2020 = period(btc_price, 2020)

nasdaq = pd.read_csv('drive/My Drive/TCC Vinicius/nasdaq-100.csv',
    ↪ on_bad_lines='skip')
nasdaq = nasdaq.set_index('date')
nasdaq.index = pd.to_datetime(nasdaq.index)
nasdaq_2017 = period(nasdaq, 2017)
nasdaq_2020 = period(nasdaq, 2020)

gold = pd.read_csv('drive/My Drive/TCC Vinicius/gold.csv')
gold = gold.rename(columns={"Date": "date"})
gold = gold.set_index('date')
gold.index = pd.to_datetime(gold.index)
gold = gold['Price'].str.replace(',', '').astype(float)
gold = gold.to_frame()
gold_2017 = period(gold, 2017)
gold_2020 = period(gold, 2020)

frac = fracDiff(btc_price, 0.4, 0.05)
adf = adfuller(frac.iloc[:,0], regression='ct')
adf_95 = adf[4]['5%']

def print_correlation():
    bit = adf_curve(btc_price)
    nas = adf_curve(nasdaq)
    gol = adf_curve(gold)

    print('BTC', bit)
    print('NASDAQ', nas)
    print('Ouro', gol)

btc_frac = fracDiff(btc_price, bit[0], 0.05)
```

```

nasdaq_frac = fracDiff(nasdaq, nas[0], 0.05)
gold_frac = fracDiff(gold, gol[0], 0.05)
print('Correlacao de Pearson BTC x NASDAQ-100 período inteiro:',
      ↪ btc_frac.corrwith(nasdaq_frac[nasdaq_frac.columns[0]]),
      ↪ interval_confidence(btc_frac, nasdaq_frac[nasdaq_frac.columns[0]]))
print('Correlacao de Pearson BTC x ouro período inteiro:',
      ↪ btc_frac.corrwith(gold_frac[gold_frac.columns[0]]),
      ↪ interval_confidence(btc_frac, gold_frac[gold_frac.columns[0]]))
print('Correlacao de Spearman BTC x NASDAQ-100 período inteiro:',
      ↪ btc_frac.corrwith(nasdaq_frac[nasdaq_frac.columns[0]], method = 'spearman'),
      ↪ interval_confidence(btc_frac, nasdaq_frac[nasdaq_frac.columns[0]], method =
      ↪ 'spearman'))
print('Correlacao de Spearman BTC x ouro período inteiro:',
      ↪ btc_frac.corrwith(gold_frac[gold_frac.columns[0]], method = 'spearman'),
      ↪ interval_confidence(btc_frac, gold_frac[gold_frac.columns[0]], method =
      ↪ 'spearman'))

print('')

bit = adf_curve(btc_price_2017)
nas = adf_curve(nasdaq_2017)
gol = adf_curve(gold_2017)

print('BTC', bit)
print('NASDAQ', nas)
print('Ouro', gol)

btc_frac = fracDiff(btc_price_2017, bit[0], 0.05)
nasdaq_frac = fracDiff(nasdaq_2017, nas[0], 0.05)
gold_frac = fracDiff(gold_2017, gol[0], 0.05)
print('Correlacao de Pearson BTC x NASDAQ-100 a partir de 2017:',
      ↪ btc_frac.corrwith(nasdaq_frac[nasdaq_frac.columns[0]]),
      ↪ interval_confidence(btc_frac, nasdaq_frac[nasdaq_frac.columns[0]]))
print('Correlacao de Pearson BTC x ouro a partir de 2017:',
      ↪ btc_frac.corrwith(gold_frac[gold_frac.columns[0]]),
      ↪ interval_confidence(btc_frac, gold_frac[gold_frac.columns[0]]))
print('Correlacao de Spearman BTC x NASDAQ-100 a partir de 2017:',
      ↪ btc_frac.corrwith(nasdaq_frac[nasdaq_frac.columns[0]], method = 'spearman'),
      ↪ interval_confidence(btc_frac, nasdaq_frac[nasdaq_frac.columns[0]], method =
      ↪ 'spearman'))
print('Correlacao de Spearman BTC x ouro a partir de 2017:',
      ↪ btc_frac.corrwith(gold_frac[gold_frac.columns[0]], method = 'spearman'),
      ↪ interval_confidence(btc_frac, gold_frac[gold_frac.columns[0]], method =
      ↪ 'spearman'))

print('')

```

```
bit = adf_curve(btc_price_2020)
nas = adf_curve(nasdaq_2020)
gol = adf_curve(gold_2020)

print('BTC', bit)
print('NASDAQ', nas)
print('Ouro', gol)

btc_frac = fracDiff(btc_price_2020, bit[0], 0.05)
nasdaq_frac = fracDiff(nasdaq_2020, nas[0], 0.05)
gold_frac = fracDiff(gold_2020, gol[0], 0.05)
print('Correlacao de Pearson BTC x NASDAQ-100 a partir de 2020:',
      ↪ btc_frac.corrwith(nasdaq_frac[nasdaq_frac.columns[0]]),
      ↪ interval_confidence(btc_frac, nasdaq_frac[nasdaq_frac.columns[0]]))
print('Correlacao de Pearson BTC x ouro a partir de 2020:',
      ↪ btc_frac.corrwith(gold_frac[gold_frac.columns[0]]),
      ↪ interval_confidence(btc_frac, gold_frac[gold_frac.columns[0]]))
print('Correlacao de Spearman BTC x NASDAQ-100 a partir de 2020:',
      ↪ btc_frac.corrwith(nasdaq_frac[nasdaq_frac.columns[0]], method = 'spearman'),
      ↪ interval_confidence(btc_frac, nasdaq_frac[nasdaq_frac.columns[0]], method =
      ↪ 'spearman'))
print('Correlacao de Spearman BTC x ouro a partir de 2020:',
      ↪ btc_frac.corrwith(gold_frac[gold_frac.columns[0]], method = 'spearman'),
      ↪ interval_confidence(btc_frac, gold_frac[gold_frac.columns[0]], method =
      ↪ 'spearman'))
return

if __name__ == '__main__':
    print_correlation()
```