Instituto de MATEMÁTICA E ESTATÍSTICA UFRGS

DEPARTAMENTO DE ESTATÍSTICA
E-mail: dest@mat.ufrgs.br

==

Trabalho de Conclusão de Curso

Arbitragem Estatística em Criptomoedas

Cristiano Sulzbach

27 de janeiro de 2018

Cristiano Sulzbach

Arbitragem Estatística em Criptomoedas

Trabalho de Conclusão apresentado à comissão de Graduação do Departamento de Estatística da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador(a): Prof. Dr. Márcio Valk

Porto Alegre
Janeiro de 2018

Cristiano Sulzbach

Arbitragem Estatística em Criptomoedas

Este Trabalho foi julgado adequado para obtenção dos créditos da disciplina Trabalho de Conclusão de Curso em Estatística e aprovado em sua forma final pela Orientador(a) e pela Banca Examinadora.

Orientador(a):_____
Prof. Dr. Márcio Valk, UFRGS
Doutor(a) pela Universidade Federal do Rio Grande do Sul, Porto Alegre, RS

Banca Examinadora:

Prof. Dr. Flávio Augusto Ziegelmann, UFRGS
Doutor(a) pela Universidade Federal do Rio Grande do Sul, Porto Alegre, RS

Porto Alegre
Janeiro de 2018

"If not Now, When?". (Autor Desconhecido)

Agradecimentos

Este trabalho representa o fim de uma jornada de estudo e dedicação. Posso dizer que a minha felicidade não está em chegar até aqui, mas sim em cada momento de aprendizagem, nas amizades criadas, e nos ótimos exemplos de professores que tive a felicidade de conviver. Quero agradecer em especial ao Prof. Dr. Hudson Torrent e Prof. Dr. Fernando Pulgatti por todos os ensinamentos e lições passadas ao longo destes 4 anos de curso, pois muito contribuíram para a minha formação. Ao meu estimado orientador, Prof. Dr. Marcio Valk, não tenho palavras para retribuir todo o apoio, empenho, sugestões e tempo dedicados para que este trabalho se torne uma realidade. Também gostaria de agradecer à minha família, que sempre apoiou as minhas decisões, compreendeu minha ausência e me deu o suporte necessário para que eu continuasse motivado. Por fim, agradeço aos meus amigos que sempre estão ao meu lado e fazem a minha vida mais feliz.

Resumo

Este trabalho propõe-se a apresentar e discutir o fenômeno mundial chamado criptomoedas. Além de explicar o conceito de *Blockchain*, que é a tecnologia por trás das moedas virtuais, vamos explorar conceitos básicos do funcionamento das criptomoedas, como a questão da segurança, transação e mineração. Além disso, vamos explorar técnicas de arbitragem estatística para criptomoedas. O desafio da aplicação de métodos estatísticos para encontrar algum tipo de padrão na cotação das criptomoedas é a alta volatilidade. Propomos uma estratégia para operar nesse mercado baseada em cointegração. A estratégia foi construída com base em 10 criptomoedas listadas na Poloniex entre o período de 5 de outubro de 2015 e 13 de setembro de 2017. Por apresentar comportamentos atípicos e de extrema volatilidade, em alguns períodos, realizou-se a verificação de quebra estrutural nas séries históricas dos preços das 10 moedas. Criou-se um ativo sintético selecionando-se ativos que cointegram e, na sequência, definiu-se a tática operacional da estratégia. Por fim, a performance da estratégia foi medida calculando-se o retorno no período analisado.

Palavras-Chave: Arbitragem, Criptomoedas, Séries Temporais, Mercado Financeiro.

Abstract

This paper proposes to present and discuss the worldwide mania called cryptocurrencies. Besides elucidate the concept of Blockchain, which is the backbone of virtual currencies, we will inquiry about basic concepts of how cryptocurrencies work, such as security, transactions and mining. Furthermore, we will explore statistical arbitrage techniques for cryptocurrencies. The challenge we face when we use statistical methods to find a pattern in cryptocurrencie's price is high volatility. We propose a strategy to operate in this market based on cointegration. The strategy was based on 10 different cryptocurrencies listed at Poloniex between October 5, 2015, and September 13, 2017. Due to the presence of atypical behavior and extreme volatility in some periods, it sought for a structural break in the historical price series of the 10 currencies. A synthetic asset was created by selecting assets that cointegrate, to define, after that, the operational tactic of the strategy. Finally, the performance of the strategy was measured by calculating the return in the examined time.

Keywords: Arbitrage, Cryptocurrency, Times Series, Financial Market.

Sumário

Lista de Figuras

Lista de Tabelas

1 Introdução

A parábola dos seis homens cegos e um elefante, que originou-se no antigo subcontinente Indiano em tempos remotos (anteriores as origens do Budismo), foi amplamente difundida até os dias de hoje. Esta parábola tem sido utilizada em diversas analogias para descrever o mercado financeiro, além de ajudar a contextualizar os objetivos cobiçados neste trabalho.

Era uma vez 6 homens cegos. Estes homens cegos queriam saber qual era a aparência de um elefante. Então eles fizeram uma viagem para a floresta e com a ajuda de seu guia encontraram um elefante domesticado. O primeiro homem cego caminhou ao lado do elefante e bateu sua cabeça nele. Ele assumiu que o elefante era como uma parede. O segundo agarrou as presas do elefante e disse que o elefante aparentava ser como uma lança. O próximo homem cego sentiu a tromba do elefante e tinha certeza que os elefantes eram parecidos com cobras. O quarto abraçou a perna do elefante e declarou que o elefante era como uma árvore. O outro pegou a orelha e disse que definitivamente elefantes eram como ventiladores. Por fim, o último cego sentiu a cauda e disse que com certeza um elefante era um tipo de corda. Dessa forma, todos os seis homens cegos perceberam um aspecto do elefante e cada um estava certo a partir de seu referencial, mas nenhum deles sabia qual era a real aparência de um elefante [Vidyamurthy, 2004].

Frequentemente, o mercado, seja o de criptomoedas ou o tradicional, é observado como um elefante. Há pessoas que dizem que previsão de mercado é como previsão de tempo, porque é possível acertar no curto prazo, mas onde o mercado estará no longo prazo é totalmente imprevisível. Outros dizem que tentar prever o mercado no curto prazo é a maneira mais rápida de fracassar, investir para o longo prazo é a filosofia deles. Alguns dirão que os mercados são eficientes, e ainda alguns outros dirão que é possível obter retornos extraordinários. Enquanto alguns seguem a análise técnica, há outros, conhecidos como fundamentalistas, que afirmam que isto é magia negra. Existem muitos modelos para avaliar o valor de um ativo, cada um sendo relevante em diferentes tempos para diferentes ativos. Teorias profundas de várias disciplinas como física, estatística, teoria gráfica, teoria dos jogos, processamento de sinais, probabilidade, e a geometria foram aplica-

das para explicar diferentes aspectos do comportamento do mercado [Vidyamurthy, 2004].

Com isso tudo dito e compreendido, este trabalho propõe-se, primeiramente, a apresentar os conceitos que norteiam o tão novo mercado das criptomoedas. Portanto, pretende-se definir o que são as criptomoedas e o Bitcoin. Também será apresentada de forma breve a tecnologia do *Blockchain* que é o que torna as criptomoedas uma realidade, além de descrever sobre ideias por trás das mesmas como descentralização, segurança, criptografia e privacidade. Conjuntamente a isto, pretende-se detalhar algumas aplicações que as criptomoedas têm. Vale mencionar que são justamente essas aplicações que criam o valor de mercado destes ativos/criptomoedas; aqui já tem-se indícios que estamos vivendo em uma nova era, na qual mais do que nunca a informação e os dados passam a ser o novo dinheiro e a moeda de troca das pessoas. Por fim, serão apresentados alguns conceitos referentes ao funcionamento, regulamentação e valorização dos ativos no mercado de criptomoedas.

Um detalhe importante é que o mercado de criptomoedas é novo e bem prematuro, apresentando diversas ineficiências até que o mesmo amadureça. Por isso, este trabalho também pretende abordar conceitos de arbitragem, que em síntese são estratégias que visam o lucro ao corrigirem estas ineficiências do mercado. Nessa abordagem, será dado um enfoque maior no entendimento de arbitragem estatística e na técnica de Pair Trading, que será a estratégia explorada neste trabalho. Assim como serão detalhados os conhecimentos estatísticos, necessários para o desenvolvimento de uma estratégia de Pair Trading baseada na abordagem de cointegração. Estão entre estes conhecimentos: o que são processos estacionários e não estacionários, como identificar um processo estacionário, como transformar um processo não estacionário em estacionário, e por fim o que é cointegração e como identificar este fenômeno.

Por último, mas não menos importante, desenvolveu-se uma estratégia de arbitragem estatística, através da ideia de Pair Trading e da utilização da abordagem de cointegração em séries temporais, que explorou as ineficiências do mercado de criptomoedas, e consequentemente obteve retornos de 80% sobre o *Bitcoin*, no período analisado.

Em resumo, como este novo mercado é muito amplo e mesmo assim já existem diversas visões sobre o mesmo, caso este trabalho consiga ampliar as perspectivas e conhecimentos da comunidade acadêmica sobre ele, através dos conceitos de arbitragem estatística e da percepção dos impactos que as tecnologias desenvolvidas em criptomoedas trazem para a sociedade, os objetivos do trabalho estarão atingidos.

2 CriptoMoedas

2.1 Descrição do Bitcoin e Criptomoedas

Durante o ano de 2017 as criptomoedas e o *Bitcoin* ganharam posição de destaque na mídia e nos grupos de conversas, onde praticamente todas as pessoas já ouviram ou leram alguma notícia sobre o assunto, além disso muitas destas pessoas têm conhecidos que já investiram em alguma criptomoeda. Apesar disto, ainda existe muita desinformação sobre o assunto e a vasta maioria da população não sabe ao certo o que é uma criptomoeda. Alguns dizem que se trata de um esquema de *ponzi*[1], outros dizem que se trata de dinheiro de *nerd*, há quem diga que se trata de uma bolha, por fim alguns consideram que as criptomoedas e o *Bitcoin* são um grande *scam*. Quando na verdade, as criptomoedas são, nada mais do que, tecnologias que permitem a transação de informações, em geral referentes a valores monetários, de maneira descentralizada e privada. Embora o conceito de criptomoeda remete ao final da década de 1980, ele popularizou-se somente quase 20 anos depois com o Bitcoin, lançado em 2009 por Satoshi Nakamoto[2], pois foi a primeira criptomoeda descentralizada bem sucedida[Farell, 2015].

Em linhas gerais, as criptomoedas são como um sistema financeiro virtual que funciona muito como uma moeda fiduciária[3], permitindo aos usuários o pagamento virtual de bens e serviços sem a necessidade de uma autoridade central. Criptomoedas dependem da transmissão de informações digitais, utilizando métodos de criptografia para garantir transações legítimas e únicas. As moedas virtuais portanto tem o poder de descentralizar o sistema financeiro e o libertar das estruturas de poder hierárquico. [Farell, 2015]

[1] Esquema de investimento fraudulento do tipo pirâmide.
[2] Não se sabe até hoje quem é Satoshi Nakamoto.
[3] Moeda que não tem lastro em nenhum metal.

2.2 Blockchain

O *Blockchain* é a principal tecnologia por traz das criptmoedas, mais especificamente é a espinha dorsal delas. Seus conceitos e principios foram introduzidos em 2008 por Satoshi Nakamoto através do *White Paper* que o mesmo publicou. Neste *White Paper* foram apresentados as ideias do *Bitcoin* e as tecnologias necessárias para o desenvolvimento deste. [Starlander, 2017].

Em primeiro lugar pode-se definir o *Blockchain* como um livro eletrônico contábil que é aberto, transparente, distribuído, online, e que qualquer pessoa com acesso à internet pode participar, a ideia é semelhante a um banco de dados relacional que pode ser compartilhado abertamente entre usuários diferentes. Basicamente, o *Blockchain* permite que os usuários criem registros imutáveis de transações, que não são armazenados em um único local, o que significa que este registro é público, pode ser verificado por qualquer pessoa, é hospedado por milhões de computadore simultaneamente e não existe uma versão centralizada desta informação que um *hacker* possa corromper. Além disso, a tecnologia do *Blockchain* traz para as criptomoedas a solução contra o $double - spend$, que é quando um usuário pode utilizar uma mesma moeda virtual inúmeras vezes. O problema do $double - spend$ será abordado mais adiante [Starlander, 2017].

Por fim, alguns aspectos gerais do *Blockchain* são: alta descentralização; identidade e segurança garantida através de criptografia; uso de uma rede P2P (ponto à ponto) para tomada de decisão; e consenso. Além disso, o fato da tecnologia ser *opensource* é totalmente favorável da perspectiva de desenvolvimento [Starlander, 2017].

Dito isto, percebe-se que, apesar de pouco conhecido, o *Blockchain* é uma tecnologia revolucionária e com capacidade de quebrar diversos paradigmas da sociedade, além de evoluir a maneira que os negócios e as indústrias operam. Enfim, pode-se enxergar o *Blockchain* como uma segunda revolução da internet, onde as pessoas não tem somente o acesso a informação, mas também podem participar ativamente das decisões [Starlander, 2017].

2.3 Descentralização

Descentralização é um conceito muito vago, pois é muito difícil definir o quão descentralizado algo é. Em relação à tecnologia do *Blockchain*, existem diversos aspectos a serem considerados. Primeiramente, quem tem o poder sobre o *Blockchain*? Segundo, quais aspectos podem ser descentralizados? Por fim, onde a descentralização é uma boa ideia? [Starlander, 2017].

O poder sobre o *Blockchain* é comumente dividido. Por exemplo, há um time de desenvolvedores que escrevem o *rulebook*[4], que a maioria dos usuários seguem ao usarem o código. Todavia, mineradores[5] também têm poder, pois os mesmos "escrevem"o histórico de transações, que é apenas consistente com as regras de consenso de mineradores. Além destes, os investidores têm ainda mais poder, pois são eles que determinam se a moeda do *Blockchain* tem algum valor, e no caso de um *hardfork*[6], são eles quem decidem qual *fork* vai vingar. Vendedores e clientes também tem poder, uma vez que são eles que geram a demanda primária e, consequentemente, o preço a longo prazo. Com isso, percebe-se que todos estes participantes são necessários para haver um ecossistema totalmente funcional. Logo, só há descentralização quando o poder é distribuído entre todos os participantes do mercado [Starlander, 2017].

Além desse conceito, a descentralização no *Blockchain* refere-se a alternativas bancárias contrárias às instituições tradicionais como o governo e instituições financeiras. Por meio dele, pode-se fazer acordos fora do "sistema legal", podendo ser intermediados por uma terceira parte confiável, que não seja uma destas instituições tradicionais. Por exemplo, há a possibilidade de ter pagamentos descentralizados e derivativos financeiros, sem a necessidade de um banco intermediando a transação e os acordos [Starlander, 2017] .

2.4 Segurança

Como o *Bitcoin*, além de outras criptomoedas, é uma moeda virtual com notável valor de mercado, não faltam motivos para que pessoas má intencionadas tentem explorar brechas de segurança para obter lucros. Além da questão do *double spending*, existem outras preocupações como um ataque de 51%, que daria poder para um indivíduo ou grupo de indivíduos controlarem toda a rede do *Bitcoin*, assim permitindo que os mesmos possam determinar quais transações são registradas, ou não, no *Blockchain*, dando abertura para fraudes na rede. Por esses motivos, percebe-se que os algoritmos de criptografia têm grande impacto para a segurança e a privacidade das implementações do Bitcoin.

2.4.1 Criptografia

O procedimento de transações de *Bitcoin* utiliza criptografia para verificar as transações, processar pagamentos, e controlar a oferta de Bitcoins. A criptografia implementada nos protocolos do *Bitcoin* não

[4]Livro de regras que o Blockchain segue

[5]Pessoas que são responsáveis por adicionar registros de transações ao Blockchain.

[6]Mudança no protocolo que faz com que existam duas cadeias de bloco, uma com a modificação do protocolo e a outra sem.

é nova. Na verdade, ela já é utilizada em uma variedade de aplicações para segurança de informação [Badev, 2014].

O *Bitcoin* utiliza basicamente de dois métodos de criptografia: assinaturas digitais e funções de criptografia *hash*. Resumidamente, o primeiro permite que as duas partes envolvidas em uma transação troquem instruções entre sí, e o segundo é utilizado para obrigar a "disciplina"no arquivamento da transação no livro contábil público [Badev, 2014].

Assinatura Digital

Assinaturas Digitais são uma maneira de autenticar uma mensagem entre um emissor e um receptor, de forma que seja garantido [Badev, 2014]:

- (I) Autenticação: O receptor pode verificar que a mensagem veio do emissor;
- (II) Não recusável: O emissor não pode negar o envio da mensagem;
- (III) Integridade: A mensagem não pode ser alterada;

A implementação de assinaturas digitais envolve a criptografia de chaves públicas, onde um par de chaves - pública e privada - são geradas com certas características desejáveis.

Funções de Criptografia Hash

Em geral, funções de criptografia *Hash* têm como entrada uma *string* de tamanho arbitrário e possuem como retorno uma *string* com tamanho pré-determinado. A mensagem de entrada aqui será tratada como m e a saída como *hash* h. A função é determinística, o que significa que a mesma entrada m irá sempre resultar na mesma saída h. Todavia, conhecer o *hash* da mensagem revela pouco sobre a mensagem. Este aspecto é fundamental para as funções de *hash* e pode ser formalmente descrito como [Badev, 2014]:

- (I) Pode receber qualquer string como entrada e produz uma saída de tamanho fixo, geralmente de 256 bites;
- (II) Precisa ser eficiente computacionalmente;
- (III) Precisa ser livre de colisões. Ninguém pode encontrar um x e y, tal que $x \neq$ y ao mesmo tempo que $H(x) = H(y)$. A colisão até pode existir, mas ela precisa ser impossível de ser encontrada computacionalmente.

Outra propriedade desejável na função de *hash* é que pequenas mudanças na mensagem m acarreta em mudanças significativas no *hash* $h = hash(m)$. Isto faz com que seja praticamente impossível alguém inferir o conteúdo da mensagem através do *hash*. Em resumo, a saída da função de *hash* deve ser muito imprevisível, mesmo que seja determinística. O *Bitcoin* utiliza o SHA-256, um tipo de algortimo *hash* desenvolvido pela *National Security Agency* e publicado pelo *National Istitute of Standards and Tecnology* [Badev, 2014].

2.4.2 Double Spending

Para entender o problema de $double - spend$, imagine que temos duas pessoas: Alice e Roberto. Alice é uma consumidora de produtos digitais vendidos online por Roberto, que provê serviços em troca de pagamentos em *Bitcoin*. O serviço de Roberto permite aos clientes que realizarem o pagamento o download de algum software de interesse. Então, como um ataque de $double - spend$ pode acontecer? Alice escolhe um produto na loja de Roberto e solicita o pagamento. Posteriormente, Alice cria uma transação de *Bitcoin* para a carteira de Roberto e transmite isso para a rede. Então, digamos que algum nó honesto criou o próximo bloco e incluiu a transação naquele bloco. Com isso, tem-se um bloco que foi criado por um nó honesto, contendo a transação que representa o pagamento realizado por Alice para Roberto [Narayanan, 2016].

Ao verificar a transação incluída no *Blockchain*, Roberto conclui que Alice realizou o pagamento e permite que ela faça o *download* do *software*. Contudo, suponha que o próximo nó aleatório selecionado seja controlado por Alice. Agora, como Alice pode propor o próximo bloco, ela pode ignorar o bloco que contém o pagamento para Roberto, ao apontar para o bloco anterior. Além disso, no bloco que ela sugerir, ela também pode incluir uma transação que transfere as moedas que ela estava enviando para Roberto para uma diferente carteira, da qual ela é dona. Este é o padrão clássico de um ataque de $double - spend$. Como as transações "gastaram"duas vezes as mesmas moedas, apenas uma delas pode ser incluída no *Blockchain*. Caso a Alice seja bem sucedida em incluir o pagamento para sua própria carteira no *Blockchain*, a transação que ela pagou Roberto não tem utilidade e nunca mais poderá ser incluída no *Blockchain* [Narayanan, 2016].

Como se pode verificar se a tentativa de $double - spend$ será bem sucedida ou não? Isto vai depender de qual dos blocos permanecerá a longo prazo no *Blockchain* - o que contém a transação de Alice para Roberto, ou o que contém a transação de Alice para Alice. O que determina qual bloco será incluído? Nós honestos seguem a política de estender a ramificação válida mais longa, então qual ramificação será estendida? Não existe uma resposta certa, pois no momento as duas ramificações tem o mesmo tamanho - elas apenas divergiram no último bloco e ambos os blocos são válidos. O nó que escolher o próximo bloco deve decidir sobre qual dos dois blocos ele quer ser construído, e esta escolha determinará se o ataque será, ou não, bem sucedido [Narayanan, 2016].

Vale ressaltar que sobre uma perspectiva moral, há uma clara diferença entre o bloco contendo a transação que paga o Roberto e aquele que Alice faz o ataque de *doublespend* com as moedas. Mas a distinção é apenas baseada no nosso conhecimento sobre a história. Contudo, sobre uma perspectiva da tecnologia, as duas transações são completamente iguais e ambos os blocos são válidos. Os nós que estão olhando para isto não tem capacidade de distinguir qual das duas transações é mais legítima [Narayanan, 2016].

2.5 Carteira de Bitcoin

Para utilizar o Bitcoin é necessário que o usuário tenha uma carteira. Esta carteira possui um par de chaves (pública e privada), que pode ser entendida como uma conta de usuário. Existe uma infinidade de carteiras disponíveis para o Bitcoin, como carteiras de *software, hardware, paper, brain e online*. As carteiras de *software* são a maneira mais comum de se utilizar o Bitcoin. Para uma carteira de *software* é necessário haver uma instância local rodando *Bitcoin*. Os desenvolvedores do *Bitcoin* implementaram um protocolo do *Bitcoin* (bitcoin.org) para realizar esta ação. Se trata de um *client* que processa todo o *Blockchain*. Já existem diversas implementações para o mesmo, como o *Armory* (bitcoinarmory.com), ou *Electrum* (electrum.org). Carteiras *onlines* como *blockchain.info* ou o *Coinbase* (coinbase.com) são alternativas populares para participar do mercado. Elas administram os recursos de forma centralizada, ou seguem uma abordagem híbrida, na qual a carteira é criptografada e a maioria das operações acontecem diretamente pelo navegador do usuário. Todas as carteiras *online* e de *software* são mais vulneráveis a problemas de segurança, uma vez que se um fraudador obtiver acesso as máquinas do usuário, é possível que ele obtenha acesso à carteira deste usuário [Tschorsch, 2015].

Em geral, uma carteira de *hardware* é aquela que usa um dispositivo separado para armazenar os *Bitcoins*, e que geralmente opera *offline*. Como o dispositivo não fica conectado diretamente com a rede, é muito mais difícil que um criminoso tenha acesso aos valores armazenados. Para obter maior segurança, também existem as carteiras *paper* e *brain*. Basicamente, uma carteira de *paper* armazena as chaves que contêm as moedas *offline*, como um documento físico. Já a carteira *brain* armazena a chave na mente do usuário pela memorização de uma frase-senha. A frase-senha é transformada em uma chave privada, da qual a chave pública e o endereço *Bitcoin* são derivados. Para evitar ataques de força bruta, a frase-senha precisa ser suficientemente longa. [Tschorsch, 2015].

2.6 Transação de Bitcoin

De forma simples e clara, uma transação diz para a rede que o dono de um número de Bitcoins autorizou a transferência de alguns desses Bitcoins para outro dono. Este novo dono agora pode gastar esses Bitcoins ao criar uma transação que autoriza a transferência para outro dono e assim por diante, em uma cadeia de donos [Antonopoulos, 2014].

Transações são como linhas em um livro de contabilidade de entrada dupla. Sendo simplista, cada transação contém uma ou mais entradas, que são debitadas de uma carteira. Do outro lado da transação, há uma ou mais saídas, nas quais o Bitcoin é creditado. As entradas e saídas

(débitos e créditos) não tem necessariamente a mesma quantidade. Ao invés disso, as saídas geralmente têm um valor um pouco menor do que as entradas, no qual esta diferença é decorrente da taxa de transação, que é um pagamento para o minerador que incluiu a transação no livro contábil. [Antonopoulos, 2014]

Além disso, os processos de transações de Bitcoin tem mecanismos que garantem que:

- (a) A verificação de cada transação é distribuída entre vários participantes da rede;
- (b) O registro de cada transação é discretizado no tempo, isto é, as transações são linearmente ordenadas em tempos consecutivos;
- (c) Os participantes da rede de pagamento devem ser recompensados pelo registro da transação;
- (d) Múltiplos nós controlam o registro de cada transação.

Na sequência, será revisado o processo de transação destacando as propriedades mencionadas.

2.6.1 Iniciando uma Transação

Suponha que Alice queira enviar 1 *Bitcoin* para Roberto utilizando a rede Bitcoin. Para fazer isso, ambos precisam term endereços Bitcoin. Digamos $address^{Alice}$ e $address^{Roberto}$. Então Alice precisa emitir e autenticar digitalmente uma mensagem como:

$$address^{Alice} \text{ está enviando } adddress^{Roberto} \text{ 1 bitcoin.}$$

Uma vez que Alice assina uma mensagem de transação, com sua chave privada e transmite, cada pessoa na rede Bitcoin pode verificar se foi Alice quem emitiu a mensagem e se a mensagem não foi adulterada. Além disso, como verificado anteriormente, as assinaturas digitais garantem que ninguém mais poderia ter assinado esta mensagem, ou seja, Alice não pode negar ter assinado [Badev, 2014].

2.6.2 Verificando uma Transação

Antes de executar uma transação, o protocolo do Bitcoin precisa verificar dois aspectos da mensagem: "$address^{Alice}$ está enviando $address^{Roberto}$ 1 bitcoin". Primeiro, foi a Alice quem transmitiu a mensagem? Como relatado anteriormente, a assinatura digital garante que apenas o dono da chave privada poderia ter assinado esta mensagem. Segundo, há fundos suficientes para garantir que a transação seja efetuada? Na sequência, é discutido como o protocolo *Bitcoin* lida com esta situação, em um cenário hipotético [Badev, 2014].

Suponha que exista um participante que mantenha registros de todos os saldos de contas e receba cada pedido de transação. Fora isto, suponha que o protocolo exija que as transações sejam aceitas sequencialmente. Por exemplo, todo dia há no máximo uma transação aceita para verificação e apuração. Seria trivial em termos de esforço para este participante verificar a integridade da transação requerida e a disponibilidade de fundos, e assim, na sequência, proceder com o registro da transação. Além disso, o fato de que as transações são aceitas sequencialmente garante que mensagens duplicadas serão rapidamente detectadas [Badev, 2014].

De forma mais geral, embora a manutenção dos registros e a verificação das transações sejam funções básicas de todos os sistemas de pagamento eletrônicos, essas funções geralmente ocorrem através de contagens privadas mantidas por terceiros. Sistemas descentralizados como o *Bitcoin* substituem os intermediários e os registros mantidos por eles, com o livro contábil público mantido por um sistema de informação distribuído. Em particular, o livro contábil público no sistema permite uma abordagem descentralizada para a verificação da transação [Badev, 2014].

2.6.3 Atualização do Blockchain

Após a verificação inicial de uma mensagem de transação assinada, um grupo de participantes da rede Bitcoin compete para registrar a transação no *Blockchain*. Primeiramente, as transações de grupo de nós concorrentes, transmitidas desde o último registro no *Blockchain*. O bloco é, então, usado para definir uma tarefa computacionalmente intensiva. O vencedor da competição é o nó que resolver primeiro a tarefa. Uma vez que o vencedor é determinado, o registro da transação é concluído. Assim, o nó vencedor tem o direito de fazer o registro e coletar a recompensa, que é um valor predeterminado de Bitcoins [Badev, 2014].

2.7 Mineração e Incentivos

Por convenção, a primeira transação de cada bloco é uma transação especial que cria um novo Bitcoin para o criador do bloco. Ao fazer isto, criam-se incentivos para que os nós deem suporte à rede, além de prover uma maneira de distribuir novas moedas para circular no mercado, uma vez que não existe uma autoridade central para emiti-las. A adição estável e constante de novas moedas é análoga à mineradores de ouro que gastam recursos para adicionar ouro em circulação. Para o Bitcoin, este recurso é o tempo de CPU e a eletricidade que é gasta [Nak, 2008].

Os incentivos também podem vir das taxas de transação. Se o montante recebido de uma transação é menor do que o valor enviado, a diferença

no preço é a taxa de transação, que é adicionada para agregar valor ao bloco que a contém. Quando um número pré-determinado de moedas estiver em circulação, o incentivo será exclusivamente baseado na taxa de transação e será livre de inflação (não haverá novas moedas entrando no mercado) [Nak, 2008].

Os incentivos ajudam os nós a se manterem honestos. Por exemplo, se uma pessoa má intencionada tiver mais poder computacional que todos os nós honestos, ela poderia escolher entre fraudar pessoas e receber seus pagamentos, ou usar esta capacidade computacional para gerar novas moedas. Esta pessoa deveria achar mais rentável seguir as regras, dado que pelas regras ela garante mais novas moedas que todos os outros juntos; ao invés de que fraudar o sistema e assim a validade de seu próprio capital [Nak, 2008].

2.8 Privacidade

O artigo inicial do *Bitcoin* descreve brevemente considerações sobre privacidade: em contraste aos bancos tradicionais - isto é, modelos de terceiros confiáveis que limitam o acesso às informações negociadas - o *Blockchain* revela publicamente todos os dados de transações. Contudo, o endereço público no *Blockchain* pretende garantir o pseudônimo, logo, a abertura do histórico de transações não implica, automaticamente, em identificação. Para garantir esta característica, um novo par de chaves (e consequentemente um novo endereço) precisa ser usado para cada transação [Tschorsch, 2015].

2.9 Aplicações

O número de possíveis aplicações da tecnologia do *Blockchain* é vasto e tem potencial para revolucionar diversas áreas da sociedade. Na sequência, serão detalhadas estas aplicações com base nas informações disponibilizadas no GitHub.

2.9.1 Contratos Inteligentes

Livros contábeis distribuídos permitem a codificação de simples contratos que são executados quando condições específicas acontecem. *Ethereum* é um projeto *Blockchain* de código aberto que foi desenvolvido especificamente para tornar isto possível. Apesar de ainda estar em etapas iniciais, o *Ethereum* tem potencial para alavancar o uso do *Blockchain* em uma escala global. Atualmente, com o nível de desenvolvimento da tecnologia, contratos inteligentes podem ser programados para realizar funções simples. Por exemplo, um derivativo poderia ser pago quando um instrumento financeiro atingir determinado patamar.

Com o uso da tecnologia *BlockChain* e do Bitcoin o pagamento poderia ser automatizado.

2.9.2 Economia Compartilhada

Com companhias como Uber e AirBnB prosperando, a economia compartilhada já tem provado seu sucesso. Atualmente, todavia, usuários que desejam utilizar serviços de transporte compartilhado precisam de um intermediário como o Uber. Ao permitir pagamentos ponto a ponto, o *Blockchain* abre as portas para interações diretas entre as duas partes, ou seja, obtêm-se uma economia compartilhada descentralizada pura. Como exemplo, OpenBazaar utiliza o *Blockchain* para criar um eBay ponto a ponto. Basta realizar o *download* do aplicativo e você pode negociar com vendedores sem a necessidade de de pagar taxas de transação.

2.9.3 Crowdfunding

Iniciativas de *crowdfunding* como Kickstart, Gofundme, Biva e Nexoos estão trabalhando para melhorar a economia ponto a ponto. A popularidade destes sites sugere que as pessoas querem ter acesso direto ao desenvolvimento de produtos. O *Blockchain* coloca este interesse em outro patamar, pois possibilita a criação de fundos de capital de risco $crowd - sourced$. Em 2016, em um experimento, o DAO (Descentralized Autonomous Organization), captou $200 milhões de de dólares em apenas 2 meses. Participantes compraram "DAO tokens", o que permitiu que eles pudessem votar em fundos de investimento de capital de risco através de contratos inteligentes (O poder de voto era proporcional ao número de DAO que cada pessoa possuía). Isto sugere que o *Blockchain* tem potencial para criar um novo paradigma de cooperação econômica.

2.9.4 Governança

Por tornar os resultados totalmente transparentes e acessíveis ao público, tecnologias de banco de dados distribuídos podem trazer transparência para eleições ou qualquer outro instrumento que necessite votação. Contratos inteligentes baseados em Ethereum ajudam a automatizar o processo. O aplicativo *Boardroom* viabiliza que organizações tenham a tomada de decisões no *Blockchain*. Na prática, isto significa que a governança da companhia torne-se totalmente transparente e verificável ao gerenciar ativos, patrimônio ou informação digital.

2.9.5 Auditoria de Cadeias de Suprimento

Consumidores, cada vez mais, querem saber aspectos éticos e da procedência dos produtos que compram. Livros contábeis distribuídos proporcionam uma maneira fácil de verificar se a "história"por trás do produto é genuína. A transparência acontece após o *BlockChain* registrar a data e a localização que corresponde a determinado produto. Ao utilizar o *Blockchain* do Ethereum, por exemplo, existe um projeto que garante que os peixes vendidos para restaurantes de Sushi no Japão advém da pesca sustentável de seus fornecedores na Indonésia.

2.9.6 Armazenamento de Dados

O armazenamento descentralizado de arquivos na internet proporciona diversos benefícios. Distribuir os dados por toda a rede os protege de serem hackeados ou perdidos. Se a internet fosse constituída apenas por sites descentralizados, poderia-se aumentar a velocidade de transferência de dados e de *streaming*. Uma melhoria assim não é apenas conveniente, mas sim uma atualização necessária para os atuais sistemas que se encontram sobrecarregados.

2.9.7 Previsão de Mercado

Já é provado que o *crowdsourcing* de predições sobre a probabilidade de algum evento acontecer tem alta precisão. Ao tirar a média de opiniões anula-se o viés de opiniões distorcidas. Já existem tecnologias que oferecem um pagamento conforme o resultado da previsão. O aplicativo de previsão de mercado Augur oferece moedas pelo resultado de eventos do mundo real. Participantes podem ganhar dinheiro ao comprar "a previsão correta". Quanto maior for a participação sobre o resultado correto, maior será o retorno. Com um pequeno investimento (menos de R$ 5), qualquer pessoa pode perguntar uma questão, criando um mercado baseado no resultado previsto, e coletar metade das taxas de transação geradas pelo mercado.

2.9.8 Proteção da Propriedade Intelectual

Sabe-se que as informações digitais podem ser infinitamente reproduzidas e distribuídas amplamente pela internet. Isto possibilitou que pessoas de todos os continentes tivessem acesso a uma mina de ouro de informações e conteúdo gratuito. Contudo, detentores de direitos autorais não tiveram tanta sorte, pois perderam controle sobre a propriedade intelectual e, assim, tiveram danos financeiros como consequência. Contratos inteligentes podem proteger direitos autorais e automatizar a venda *online* de trabalhos criativos, eliminando o risco da cópia e distribuição de arquivos. Mycelia, por exemplo, utiliza o *Blockchain* para criar um sistema distribuído de música ponto a ponto. Mycelia

permite que os artistas vendam as músicas diretamente para o público, ao utilizar contratos inteligentes para automatizar o processo.

2.9.9 Internet das Coisas

A Internet das Coisas (IoT) refere-se à gestão controlada da rede por certos tipos de dispositivos eletrônicos. Por exemplo, o monitoramento da temperatura do ar em um armazenamento poderia ser automatizado ao utilizar contratos inteligentes. Ao utilizá-los torna-se possível a automação de gerenciamento desses sistemas remotos. Uma combinação de *software*, sensores, e rede facilitam a troca de informações entre objetos e mecanismos. O resultado é que há um ganho na eficiência do sistema e uma melhora nos custos de monitoramento.

Os gigantes da indústria de manufatura, tecnologia e telecomunicações estão lutando pela dominância da IoT. As aplicações da Internet das Coisas envolvem a manutenção preditiva de peças mecânicas; análise de dados e gerenciamento de sistemas automatizados em grande escala.

2.9.10 Microgrids de Vizinhança

A tecnologia de *Blockchain* permite comprar e vender energia renovável gerada por *microgrids* de vizinhança. Quando um painel solar gera energia excedente, contratos inteligentes automaticamente a redistribuem. Outros tipos de automatização de contratos inteligentes terão outras aplicações assim que a Internet das Coisas se tornar realidade. *Consensys* é uma companhia que está trabalhando em um protótipo para automatizar a redistribuição de energia gerada por *microgrids*, através do uso de contratos inteligentes do *Ethereum*. Esta ideia também mostra uma das funcionalidades da Internet das Coisas.

2.9.11 Gerenciamento de Dados

Atualmente, em troca de seus dados, as pessoas podem navegar em plataformas de mídia social gratuitamente. No futuro, os usuários terão a habilidade de gerenciar e vender os dados que suas atividades *online* criaram. Como o *Bitcoin*, ou outra criptomoeda, pode ser facilmente distribuído em pequenas quantidades, ele será a moeda utilizada para realizar este tipo de transação.

2.9.12 Registro de Propriedade

Como um livro contábil público, o *Blockchain* pode facilmente armazenar todos os tipos de informações com maior eficiência. Títulos de propriedade são um exemplo disso. Eles geralmente tendem a ser suscetíveis a fraudes, como também existem diversos custos de trabalho para administrá-los.

2.9.13 Negociação de Ações

O potencial de maior eficiência na liquidação de ações é um forte caso de uso para o *Blockchain* na negociação de ações. Quando executadas ponto a ponto, as confirmações do mercado tornam-se quase instantâneas (em oposição a três dias no atual sistema). Potencialmente, isso significa que os intermediários - auditores e agentes de custódia - são removidos do processo.

Numerosas bolsas de ações e commodities estão criando protótipos de aplicativos de *Blockchain* para os serviços por eles oferecidos, incluindo a ASX (*Australian Securities Exchange*), a *Deutsche Börse* (Bolsa de Valores de Frankfurt) e o JPX (*Japan Exchange Group*).

2.10 Mercado

2.10.1 Regulamentação

Embora o mercado de criptomoedas tenha potencial para revolucionar a maneira como se lida com o dinheiro e diversas outras tecnologias, a sua introdução em uma escala global é repleta de desafios e possíveis armadilhas, já que as criptomoedas não são universalmente reconhecidas como meios oficiais de pagamento [Farell, 2015].

Os sistemas regulatórios estão crescendo, com inúmeras abordagens sendo tomadas por vários governos. As medidas regulatórias atuais estão em sua infância e continuam a evoluir com a indústria em rápida expansão. Os regulamentos oferecerão maior legitimidade para as moedas obterem aceitação em massa, sendo padronizados os elementos do mercado, diminuindo um pouco a volatilidade. O objetivo final é o mesmo: reduzir a fraude, proteger os consumidores, respeitar as sanções econômicas e instituir métodos de tributação [Farell, 2015].

Contudo, deve-se observar que a regulamentação do Bitcoin e de outras criptomoedas vai contra seus princípios, uma vez que o atual sistema financeiro baseado na confiança nos agentes econômicos é muito frágil, expondo a sociedade como um todo aos riscos inerentes a esse ambiente. Assim, as criptomoedas foram criadas para alcançar um objetivo muito claro: promover a descentralização na execução de transações comerciais na Internet (implementação de uma liberdade monetária), isto é, possibilitar aos vendedores e compradores a realização de suas transações independentemente, sem a necessidade de instituições financeiras intermediadoras [Silva].

2.10.2 Corretoras

As corretoras de criptomoedas são extremamente importantes. As decisões, ações e erros cometidos por essas empresas afetam a percepção

pública sobre as criptomoedas. Para que uma moeda cresça, as pessoas precisam poder comprar e vender com facilidade e segurança, e as corretoras são o que torna isso possível.

Somente agora as criptomoedas estão se tornando *mainstream*. À medida que elas amadurecem como ativos financeiros, as corretoras competem para tornarem-se a plataforma preferida para investidores. Com as oportunidades e os desafios futuros no mundo de criptomoedas, as corretoras desempenharão um papel fundamental na moldagem da inovação, aceitação e tendências futuras neste campo.

Dentre as principais corretoras para negociar criptomoedas tem-se:

- Bitcointoyou (Brasileira);
- Bitfinex;
- Bittrex ;
- CEX;
- Coinbase;
- Cryptopia;
- Binance;
- Foxbit (Brasileira);
- Poloniex;
- Kraken;

2.10.3 Valorização

Na tabela a seguir é apresentada a valorização das 26 principais criptomoedas entre as datas de 1º de janeiro de 2017 a 22 de setembro de 2017. Como pode-se observar, o ganho das criptomoedas foi altíssimo, muito acima de qualquer outro investimento disponível no mercado. Isso comprova a alta demanda e a aceitação popular das criptomoedas.

Tabela 2.1: Valorização das principais Criptomoedas

Moeda	Preço	Retorno (Ano)
Bitcoin (BTC)	$ 3,643.07	265%
Ethereum (ETH)	$ 259.23	3114%
Bitcoin Cash (BCH)	$ 420.04	0%
Ripple (XRP)	$ 0.17	2590%
Dash (DASH)	$ 335.60	2846%
Litecoin (LTC)	$ 46.60	935%
Nem (XEM)	$ 0.21	6103%
Iota (IOT)	$ 0.50	-4%
Monero (XMR)	$ 86.43	535%
Ethereum Classic (ETC)	$ 10.09	621%
Neo (NEO)	$ 17.50	12240%
OmiseGo (OMG)	$ 8.68	630%
BitConnect (BCC)	$ 109.93	75025%
Lisk (LSK)	$ 5.40	3591%
Tether (USDT)	$ 1.00	0%
Qtum (QTUM)	$ 7.71	-34%
Stratis (STRAT)	$ 4.04	5202%
Zcash (ZEC)	$ 174.99	261%
Waves (WAVES)	$ 3.64	1543%
Ark (ARK)	$ 2.83	8235%
Eos (EOS)	$ 0.59	-73%
Golem (GNT)	$ 0.24	2362%
BitShares (BTS)	$ 0.07	1650%
Siacoin (SC)	$ 0.00	1630%
Dogecoin (DOGE)	$ 0.00	237%
Stratus (SNT)	$ 0.02	-61%

Fonte: CoinGecko

2.10.4 MarketCap e Liquidez

Um dos pontos importantes a ser observado no mercado de Criptomoedas é que a capitalização de mercado cresceu excepcionalmente durante o ano de 2017. Pode-se observar, na figura 2.1, que em Fevereiro de 2017 a capitalização de mercado, somada das 26 maiores criptomoedas, era de aproximadamente $ 20 bilhões de dólares. Já em Setembro, esta capitalização chegou a atingir valores superiores a $ 160 bilhões de dólares, superando o valor de diversas empresas consagradas, como a Petrobras e a Vale.

Também deve-se analisar o volume de negociação diário típico das principais criptomoedas combinadas atualmente é superior a $ 2 bilhões de dólares. Alguns dias com valores superiores até $ 6 bilhões. Embora isso não signifique, necessariamente, que as pessoas estão gastando estes valores, mostra que as criptomoedas estão sendo usadas cada vez

mais frequentemente para qualquer finalidade que o usuário pretenda. Isso, por si só, é um grande passo, pois mostra que há um amadurecimento das criptomoedas.

Figura 2.1: MarketCap vs Volume de Tx das 26 maiores Criptomoedas
Fonte: CoinGecko

2.10.5 Initial Coin Offerings

Initial coin offerings (ICOs) é uma nova forma de *crowdfunding*. Ao contrário dos tradicionais *crowdfunds*, ICOs ignoram a rigorosidade e a regulação requerida no processo de captação de capital por bancos e *venture capitalists*. Atualmente, o uso de ICO's é limitado ao lançamento de novas criptomoedas no mercado [Starlander, 2017].

Em uma campanha de ICO, uma porcentagem da criptomoeda é vendida para financiadores do projeto em troca de outras criptomoedas estabelecidas no mercado, como o Bitcoin ou Ethereum, e, em alguns casos, por moedas fiduciárias. Como resultado, a companhia obtém o capital para financiar o desenvolvimento do produto e os membros ganham uma participação dos tokens da criptomoeda. Estes ICOs vêm se mostrando muito eficientes para iniciar projetos de criptomoedas e satisfazer as demandas do mercado [Starlander, 2017].

Existem algumas semelhanças entre ICOs e os mais tradicionais *Inicial Public Offerings* (IPOs). Todavia, existem algumas características que diferenciam os mesmos. As ações de uma companhia, distribuídas durante um IPO, sempre indicam uma parcela de propriedade da empresa, o que não é uma verdade para um ICO, no qual os tokens são na verdade unidades de uma moeda. Outra diferença fundamental se encontra no fato de que IPOs são fortemente regulados pelo governo. Isto requer que a empresa tenha diversos trabalhos burocráticos antes de inaugurar suas ações. Já para lançar um ICO, poucos requerimentos são necessários, o que significa que qualquer projeto pode lançar um ICO em pouco tempo e com pouca preparação, e qualquer pessoa pode investir no projeto, não fazendo diferença o lugar do mundo em que ela se encontra [Starlander, 2017].

Vale lembrar que o lucro não é garantido ao investir em um ICO, assim como em um IPO. As campanhas podem falhar, e, neste caso, parte das contribuições podem ser devolvidas aos investidores. Além disso, mesmo que ela suceda, não há garantia que o preço do token subirá, ou que o projeto entregará o produto prometido. Este é um risco que todos os participantes precisam estar cientes antes de decidir investir em um ICO [Starlander, 2017].

3 Arbitragem Estatística

3.1 Visão Geral: Arbitragem

Para entender o conceito de arbitragem no mundo financeiro, é preciso formalizar a ideia de mercado eficiente. Conforme apresentado no artigo [Junior, 2004] mercado eficiente é aquele em que o preço dos ativos negociados sempre reflete inteiramente as informações disponíveis sobre o mesmo. Com isso, fica clara a ideia de que só é possível obter retornos maiores assumindo riscos maiores. Ou seja, assume-se que os investidores são racionais e têm todas as informações necessárias para avaliar determinado ativo, e, portanto, pagam o preço "justo"por este ativo; e que todas as operações irracionais, que não refletem o preço praticado pelo mercado, sejam aleatórias, cancelando-se e não criando uma distorção no mercado.

Todavia, existem diversas anomalias como o efeito de volatilidade, de momentum, de valor e de tamanho que põem em xeque esta hipótese de mercado eficiente. Basicamente, estas anomalias mostram evidências de que o modelo de apreçamento não consegue explicar os retornos de certas estratégias ou carteiras de ativos. Desta maneira, estas anomalias criam um ambiente propício para o investidor utilizar mecanismos de correção das ineficiências do mercado, obtendo retornos sem risco (ou quase sem). Sabendo disso, pode-se definir arbitragem como sendo qualquer mecanismo utilizado para corrigir o preço de determinado ativo. Para contextualizar melhor este conceito, serão relatados alguns tipos de arbitragem utilizados no mercado.

3.1.1 Arbitragem Cash and Carry

A arbitragem *Cash and Carry* consegue ilustrar bem o conceito de arbitragem, pois consiste em identificar discrepâncias entre os valores praticados nas taxas de juros de empréstimos bancários e os preços que possíveis tomadores de crédito estão dispostos a pagar. Assim, o arbitrador tende a atuar como um facilitador da operação, adquirindo o empréstimo diretamente no banco e posteriormente repassando os valores, acrescidos de uma taxa de conveniência para o tomador da

operação. Por exemplo, considere que um Banco A pratica empréstimos a uma taxa de 5% ao mês e que um tomador B precisa de um empréstimo urgente e está disposto a pagar 6% ao mês. O arbitrador obterá o empréstimo no banco A e disponibilizará o mesmo para o tomador B. O seu lucro será a diferença entre as taxas negociadas, ou seja, espera-se que ele obtenha 1% de retorno a cada mês.

Entretanto, como já diz o ditado: nem tudo são flores. Na prática, existem poucas oportunidades suficientemente rentáveis para investidores obterem lucros consistemente com esta técnica.

3.1.2 Arbitragem Inter-Mercados

Como citado anteriormente, existem diversas anomalias no mercado, que por sua vez não permitem a hipótese de mercado eficiente. Uma das principais razões para elas acontecerem está associada com a liquidez dos ativos, dando margem para a arbitragem Inter-Mercados, por exemplo. A arbitragem Inter-Mercados é quando o preço de um mesmo ativo difere em diferentes locais de compra do mesmo.

Para contextualizar esta situação, pode-se pensar em duas casas de câmbio que negociam o par de moedas Real/Dólar. Digamos que a casa de câmbio A teve forte demanda de venda do dólar na última semana. Já a casa B teve alta procura para comprar dólar. Consequentemente, ambas as casas podem enfrentar empecilhos quanto à liquidez de suas moedas, na qual as casas A e B, respectivamente, dispõem de uma quantidade baixa de dólares e reais. Deste modo, ambas as corretoras de câmbio estão expostas ao risco de ficarem sem moeda e assim deixarem de operar. Para evitar que isto aconteça, a agência A passa a negociar o par Real/Dólar a um preço premium de R$ 3.18, ao passo que a agência B negocia o par de moedas ao valor de R$ 3.16.

Por conseguinte, o arbitrador perspicaz, que percebe a oportunidade, passa a comprar dólar na casa B para então revender na casa A. Um ponto importante de se mencionar é que a medida em que ocorre a atuação do arbitrador, a operação que a originou tende a desaparecer. Nesse caso, por exemplo, conforme o arbitrador comprar dólares na corretora B, a carência de reais tende a diminuir, elevando o preço a R$ 3.17; assim como quando ele vende estes mesmos dólares para a agência A, diminuindo o preço a R$ 3.17. Com isso, os valores praticados pelo mercado passam a estar alinhados novamente, pois a oferta e a demanda em ambas as casas de câmbio foram reguladas. Ou seja, tem-se uma situação que ambos os lados ganham, o arbitrador obtem seu lucro e as corretoras pudem equilibrar a liquidez de moedas.

3.1.3 Arbitragem de Fusões

Seguidamente observam-se compras e fusões de empresa no mercado. Contudo, muitas dessas aquisições não refletem o preço que vinha sendo

praticado anteriormente, ou algum acionista majoritário decide liquidar sua posição na transição, criando uma pressão vendedora e consequentemente baixando os preços. Assim, os arbitradores conseguem adquirir estas ações com deságio.

3.1.4 Arbitragem de Índice

No mercado financeiro, um índice é apenas uma cesta reunindo diversos ativos. Não obstante, em alguns casos o índice não reflete inteiramente os valores praticados pela compra individual de cada um dos ativos.

3.2 Origens da Arbitragem Estatística

Após estabelecer a ideia primordial de arbitragem, na qual ela atua como reguladora do mercado, será apresentado um resumo da história da Arbitragem Estatística.

O início da Arbitragem Estatística aponta para uma técnica de *trading* conhecida como ***pairs trading***. Após mais de 30 anos desde seu nascimento, esta técnica, que explora discrepâncias e correlações entre pares de ativos, continua sendo a essência da Arbitragem Estatística. É claro que a arbitragem estatística não é apenas uma única estratégia, mas, sim, um conjunto de estratégias que se utiliza de modelos sofisticados de estatística e matemática para analisar diferenças e padrões de preços, no intuito de obter retornos acima de média. Os conceitos matemáticos utilizados nas técnicas de Arbitragem Estatística envolvem desde séries temporais, análise de componentes principais, cointegração, redes neurais, matriz de covariância, análise de fronteira eficiente, até conceitos mais avançados da física como a minimização de energia [Latte, 2011].

Segundo a literatura [Vidyamurthy, 2004], acredita-se que a origem da técnica de ***pair trading*** originou-se do trabalho de Nunzio Tartaglia, um brilhante analista de Wall Street. Na década de 1980, Tartaglia, que trabalhava para o banco Morgan Stanley, teria reunido um grupo ultra secreto de matemáticos, físicos e cientistas da computação para desenvolver estratégias quantitativas de arbitragem que utilizassem técnicas avançadas de estatística e tecnologia da informação. As estratégias desenvolvidas pelo grupo foram automatizadas até um ponto em que elas podiam ser executadas sem interferência humana, o que era extraordinário para os padrões da época. Este grupo é conhecido por ter resultados fantásticos na casa dos 50 milhões de dólares no ano de 1987. Porém, o grupo se desfez em 1989. A partir de então, os membros se espalharam por outras instituições, bancos e para os tão famosos *Hegde Funds*, acabando por disseminar o conhecimento e aumentar a popularidade da técnica.

3.3 Pair Trading

De acordo com [Gundersen, 2014] o conceito básico de *Pair Trading* é bem simples; deve-se identificar um par de ativos com histórico de co-movimentos em seus preços. Posteriormente, quando ocorrerem desvios significantes da relação de estabilidade dos preços observados, uma posição é aberta. Esta posição consiste em, simultaneamente, comprar o ativo desvalorizado (*long position*) e a venda a descoberto do ativo sobrevalorizado (*short position*). A estratégia baseia-se no princípio que os preços relativos em um mercado estão em equilíbrio, tendo desvios deste equilíbrio possivelmente corrigidos. Logo, investir na estratégia de *pair trading* é tentar lucrar com os desvios temporários do equilíbrio.

Para entender o equilíbrio pode-se imaginar a relação entre ações ordinárias e preferenciais de um mesmo ativo, um par de ativos de um mesmo segmento do mercado como a Coca-Cola e Pepsi, ou até mesmo a relação de dois ativos que aparentemente não possuem características em comum. A fim de ilustrar os exemplos, observe os preços dos ativos da Vale e do Bradespar na figura 3.1, que trazem a relação de equilíbrio descrita. Ou seja, qualquer distorção nesta relação abrirá uma oportunidade de atuação sobre o mercado.

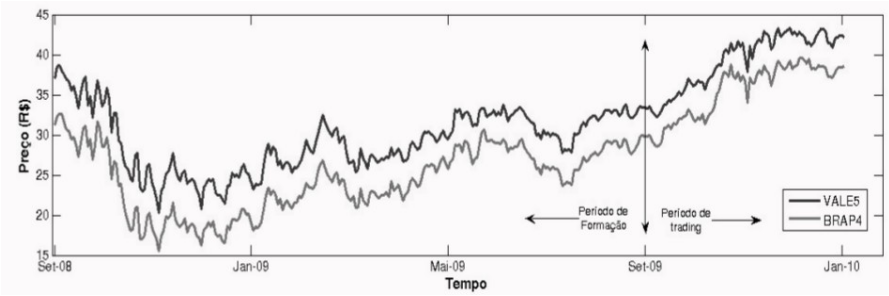


Figura 3.1: Séries de preços de VALE5 e BRAP4
Fonte: [Caldeira, 2013]

Também é apresentada uma relação entre os spreads padronizados dos preços dos dois ativos. Como percebe-se na figura 3.2, o spread é centrado em 0 e aparentemente tem comportamento aleatório e estacionário (a definição de processo estacionário é abordada posteriormente neste trabalho), o que permite estabelecer pontos de compra e venda bem definidos dos ativos. Uma vez que o spread atingir a linha inferior no gráfico deve-se comprar as ações da Vale e vender as da Bradespar. Quando o spread padronizado entre elas retornar a 0, deve-se liquidar a posição. Caso contrário, ou seja, se o spread atingir a linha superior, deve-se realizar a operação inversa: comprar Bradespar e vender Vale.

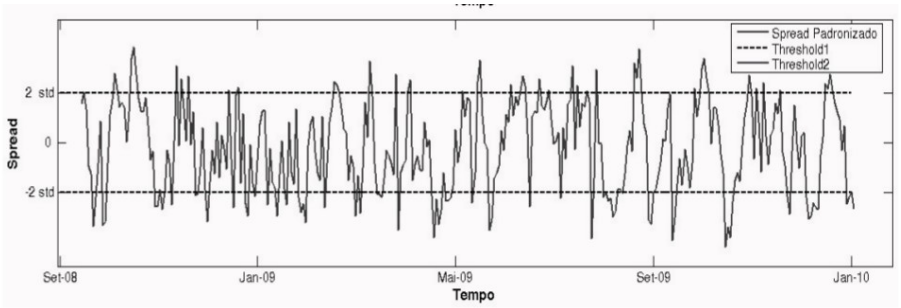


Figura 3.2: Spread Padronizado VALE5 e BRAP4
Fonte: [Caldeira, 2013]

3.4 Elementos da Arbitragem Estatística

Para [Parreira, 2007] existem 3 componentes básicos para a construção de uma estratégia de arbitragem estatística. São eles:

- (I) Construção de relações estatísticas de preços em equilíbrio entre os ativos, de maneira que os desvios tenham um componente potencialmente previsível (através das séries temporais dos preços dos ativos).
- (II) Identificação das oportunidades de arbitragem (através da previsão das variações nas combinações apropriadas dos ativos).
- (III) Implementação das operações apropriadas (ao comprar o ativo - ou conjunto de ativos - que se prevê estar subvalorizado, e vender o ativo - ou conjunto - que se prevê estar sobrevalorizado).

3.5 Considerações Teóricas

Segundo [Parreira, 2007] arbitragem estatística é uma estratégia de *trading* com custo inicial zero, auto-financiada (x(t):t>0) com valor acumulado descontado v(t) tal que:

- (I) $v(0) = 0$
- (II) $\lim_{t\to\infty} E^p[v(t)] > 0$
- (III) $\lim_{t\to\infty} P(v(t) < 0) = 0$
- (IV) $\lim_{t\to\infty} \frac{Var^P[v(t)]}{t} = 0$ se $P(v(t) < 0) > 0, \forall t < \infty$

Portanto, a estratégia: (i) é autofinanciada, ou seja, não há custos para montar uma estratégia de arbitragem; (ii) tem expectativa de lucros positivos; (iii) tem probabilidade de perda que converge para zero e (iv) a média temporal da variância converge para zero se a probabilidade de perda não se tornar zero num tempo finito [Parreira, 2007].

3.6 Medida de Performance

O índice que vamos apresentar para medir a performance é o índice Sharpe. Este índice é uma medida para calcular o retorno ajustado ao risco, logo, quanto maior este valor, melhor será. O índice tornou-se amplamente usado em todo o mundo financeiro e foi desenvolvido por William F. Sharpe.

O índice mede o excesso de retorno médio ganho em relação à uma taxa livre de risco por unidade de volatilidade. Numericamente temos que:

$$\widehat{SR_k} = \frac{\hat{\mu_k} - R_f}{\hat{\sigma_k}}$$

Onde $\hat{\mu_k}$ é o retorno médio do ativo, $\hat{\sigma_k}$ é a variância do ativo, e R_f é o retorno livre de risco do mercado (em geral é utilizado 100% do CDI como retorno livre de risco).

4 Aspectos Estatísticos

Neste capítulo são apresentados os conhecimentos estatísticos necessários para construir uma estratégia de *Pair Trading* como a estacionariedade e cointegração.

4.1 Processos Estacionários

Ao trabalhar com Séries Temporais, é importante entender o conceito de estacionariedade, pois segundo [Bueno, 2011], é através da constatação de estacionariedade que se pode proceder com inferências estatísticas sobre os parâmetros estimados.

Um processo é dito estacionário quando as distribuições de probabilidade têm estabilidade ao longo do tempo. Logo, se $\{y_t\}_{t\in\mathbb{Z}}$ é um processo estacionário, as características de y_t serão as mesmas que y_{t+k}, para qualquer valor de k. Ou seja, quando uma série se desenvolve no entorno de uma média e outras características como variância constantes ao longo do tempo, refletindo alguma forma de equilíbrio estatístico estável, a série é dita estacionária [Pereira, 2013].

Pode-se caracterizar a estacionariedade de um processo de duas formas diferentes, através da estacionariedade fraca ou estrita.

Definição: O processo estocástico ou série temporal, $\{y_t\}_{t\in\mathbb{Z}}$ é fracamente estacionário se:

1. $E|y_t|^2 < \infty$;
2. $E(y_t) = \mu$, para todo $t \in Z$; e
3. $E(y_t - \mu)(y_{t-j} - \mu) = \gamma_j$.

A primeira condição afirma apenas que o segundo momento não centrado deve ser finito, ainda que desigual em diferentes períodos. A segunda condição assegura que a média é igual para todo o período, mesmo que a distribuição da variável aleatória vá se alterando ao longo do tempo. A terceira condição estabelece que a variância é sempre igual para todo o período e que a autocovariância não depende do tempo, mas da distância temporal entre observações [Bueno, 2011].

Além desta definição, há também a estacionariedade estrita.

Definição: O processo estocástico ou série temporal, $\{y_t\}_{t\in\mathbb{Z}}$ é estritamente estacionário se a função de distribuição conjunta $\{y_{t_i}\}_{i=1}^{k}$ for igual à função de distribuição conjunta de $\{y_{t_i+h}\}_{i=1}^{k}$, $h \in Z$, isto é:

$$F(y_{t_1}, y_{t_2}, ..., y_{t_k}) = F(y_{t_1+h}, y_{t_2+h}, ..., y_{t_k+h}).$$

Ao analisar a fórmula descrita acima, tem-se que estacionariedade estrita significa que independente do tempo t analisado, os momentos populacionais devem ter as mesmas propriedades estatísticas [Bueno, 2011].

4.2 Função de Autocorrelação

Basicamente, a função de autocorrelação é o gráfico da autocorrelação contra a defasagem no tempo, ou seja, indica o quanto um processo y_t é correlacionado com ele mesmo em instantes de tempos diferentes, y_t, y_{t-1}, y_{t-2} ... [Pereira, 2013].

A partir da análise da função de autocorrelação é possível determinar alguns padrões da série, o que, por sua vez, ajuda a conhecer as características de alguma série. Esses padrões, todavia, nem sempre são claros, por serem estocásticos. Uma mesma série pode sugerir padrões alternativos de modelagem, em razão da dificuldade de inferir qual é o padrão gerador daquela função [Bueno, 2011].

Na prática, o que importa é extrair o máximo de informações da série a partir da função de autocorrelação para então gerar a melhor modelagem possível a partir dos dados observados, a ponto de gerar resíduos que sejam estatisticamente um ruído branco. Supostamente, esta modelagem levaria às melhores previsões estatísticas [Bueno, 2011].

Além disso, quando o interesse é determinar a correlação pura entre dois instantes de tempo, é possível filtrar as correlações, de forma a manter-se apenas a correlação pura entre duas observações. Este processo de filtragem implica gerar a função de autocorrelação parcial, FACP, pela qual eliminam-se as correlações implícitas entre duas observações [Bueno, 2011].

4.3 Processos Não Estacionários

Se por um lado, quando um processo se desenvolve ao longo de uma média, ele é dito estacionário, quando um processo apresenta média variando ao longo do tempo, ele é dito não estacionário. Pelo fato deste processo não apresentar média e variância constante, a estimação dos parâmetros do processo é complexa. Para contornar esta situação, uma das alternativas é diferenciar o processo até que o mesmo seja estacionário.

Os resultados obtidos ao utilizar séries temporais não estacionárias podem ser espúrios, no sentido que eles podem indicar relação entre variáveis, que na verdade não existem. Com isso, para que o processo apresente resultados consistentes e confiáveis, é necessário transformar os dados em estacionários [Pereira, 2013].

Todavia, quando se fala em transformação de dados não estacionários para estacionários, é importante lembrar que existem diferentes tipos de processos não estacionários. Como exemplos desses processos temos o passeio aleatório com e sem constante; e a tendência estacionária e estocástica. Seguem as definições dos mesmos.

4.3.1 Passeio Aleatório sem Constante

O processo pode ser definido como:

$$y_t = y_{t-1} + \epsilon_t.$$

No passeio aleatório, o valor de y em um tempo "t"deve ser igual ao último período acrescido de um componente estocástico que seja um ruído branco, onde em geral ϵ_t é independente e identicamente distribuído com média 0 e variância σ^2. O passeio aleatório é um processo de não reversão à média, ou seja, ele pode se distanciar da média para cima ou para baixo. Outra característica deste processo é que a variância aumenta ao longo do tempo e vai para o infinito a medida que o tempo vai ao infinito.

4.3.2 Passeio Aleatório com Constante

O processo pode ser definido como:

$$y_t = \alpha + y_{t-1} + \epsilon_t.$$

Se o modelo para o passeio aleatório prediz que o valor de y em um tempo "t"é igual ao último período acrescido de uma constante e um ruído branco, então o processo é um passeio aleatório com constante. Este processo também não apresenta características de retorno à média no longo prazo e tem variância dependente do tempo.

4.3.3 Tendência Determinística

O processo pode ser definido como:

$$y_t = \alpha + \beta t + \epsilon_t.$$

Com certa frequência, o passeio aleatório com constante é confundido com a tendência determinística. Ambos incluem constante e um ruído

branco, mas o valor em um tempo "t" no caso do passeio aleatório é regredido sobre o valor do último período, enquanto no caso da tendência determinística é regredido sobre uma tendência de tempo. Um processo não estacionário com tendência determinística tem média que cresce ao redor de uma tendência fixa.

4.3.4 Passeio Aleatório com Constante e Tendência Determinística

O processo pode ser definido como:

$$y_t = \alpha + y_{t-1} + \beta t + \epsilon_t.$$

É um processo não estacionário que combina passeio aleatório com constante e uma tendência determinística. Ele especifica o valor em um tempo "t"como o valor do último período, uma constante, a tendência e um componente estocástico. Isto implica que para estacionarizar o processo não é necessário somente diferenciar o mesmo, mas também realizar uma remoção da tendência.

4.3.5 Estacionarizar o Processo

Ao utilizar uma série temporal não estacionária, os modelos podem produzir resultados não fidedignos e espúrios, que levam à previsões com pouco sentido para o mundo real. Um das soluções para este problema é transformar a série em estacionária para então fazer a modelagem desta. A seguir são apresentadas as maneiras para obter tal resultado.

Um passeio aleatório com ou sem constante pode ser transformado em processo estacionário através da diferenciação correspondente de $y_t - y_{t-1} = \epsilon_t$ ou $y_t - y_{t-1} = \alpha + \epsilon_t$ e assim o processo se torna estacionário.

Já um processo com uma tendência determinística se torna estacionário após a remoção da tendência. Por exemplo, $y_t = \alpha + \beta t + \epsilon_t$ é transformado em um processo estacionário quanto subtraímos a tendência βt, matematicamente temos que:

$$y_t - \beta t = \alpha + \epsilon_t.$$

Por fim, caso seja um passeio aleatório com constante e uma tendência determinística, a remoção da tendência elimina a tendência determinística e a constante, mas a variância irá continuar indo ao infinito. Portanto, deve-se aplicar a diferenciação para remover a tendência estocástica.

4.4 Teste de Raiz Unitária

Para determinar se uma série é estacionária ou não, pode-se realizar o teste de raiz unitária. O primeiro teste para tal foi proposto por [Dickey, 1979].

Considere o seguinte modelo:

$$y_t = \phi y_{t-1} + \epsilon_t.$$

A ideia inicial é estimar esse modelo e usar o *teste* t sobre ϕ, tendo como hipótese nula $H_0 : \phi = 1$. Em geral, os pacotes econométricos reportam os testes nos coeficientes contra a hipótese nula de serem iguais a zero. Então, pode-se alterar o teste subtraindo y_{t-1} de ambos os lados:

$$\Delta y_t = (\phi - 1)y_{t-1} + \epsilon_t = \alpha y_{t-1} + \epsilon_t$$

em que se define $\alpha = \phi - 1$

Assim, $H_0 : \phi = 1$ é equivalente a $H_0 : \alpha = 0$. Por fim, se $\phi = 1$, ou seja, se rejeitarmos a hipótese nula, o processo será estacionário, do contrário será não estacionário.

4.5 Cointegração

Anteriormente, se discutiu brevemente como podemos transformar um processo não estacionário em estacionário. Assim, quando se analisa séries temporais multivariadas, onde cada uma das séries é não estacionária, há sentido em estacionarizar cada uma delas, para então analisar as mesmas. Contudo, esta não é, necessariamente, a única maneira de trabalhar com estas séries [Vidyamurthy, 2004].

Ao estudar diversas séries multivariadas para determinar estatisticamente se há uma relação de causa e efeito entre as variáveis representadas pelas séries temporais, os economistas Engle e Granger observaram o seguinte fenômeno. Mesmo que duas séries sejam não estacionárias é possível que exista uma combinação linear específica delas que seja estacionária. Este fenômeno passou a ser chamado de cointegração [Vidyamurthy, 2004].

Formalmente o conceito de cointegração definido segundo [Engle, 1987] é que se cada elemento de um vetor de uma série temporal multivariada Y_t tem mesma ordem de integração e ao mesmo tempo existe uma combinação linear de $\alpha' Y_t$ que é estacionária, é dito que esta série é cointegrável com o vetor α. Assim, assumindo $\alpha' Y_t = 0$ como um equilíbrio de longo prazo, a cointegração implica que os desvios deste equilíbrio são estacionários, com variância finita, mesmo que as séries originais sejam não estacionárias e com variância infinita.

Definição: Os componentes do vetor Y_t são ditos cointegráveis de ordem d,b, denotados Y_t $CI(d,b)$, se:

i) Todos os componentes de Y_t são I(d)

ii) Existe um vetor $\alpha(\neq 0)$, tal que $Z_t = \alpha' Y_t$ $I(d-b)$, $b > 0$. Este vetor é cointegrável.

Imaginando que d = 1, b = 1, a cointegração significaria que se os componentes de Y_t são todos I(1), então o erro de equilíbrio seria I(0). Logo Z_t deve mover-se na vizinhança de 0 (considerando que Z_t tenha média 0). Em outras palavras, o equilíbrio ocasionalmente deve ocorrer, ou pelo menos deve se aproximar desta relação [Bueno, 2011].

4.6 Cointegração Múltipla

Em modelos com mais de duas variáveis, nem sempre é imediato concluir se as séries cointegram. Se duas variáveis têm ordem de integração diferentes, qualquer combinação linear entre elas resultará em uma variável cuja ordem de integração será a de maior ordem. Ou seja, a ordem de integração da variável de maior ordem domina a da variável de menor ordem. Esse fato sugere a necessidade das variáveis serem de mesma ordem para haver cointegração. Contudo, em um modelo em que o número de variáveis endógenas é maior que 2, nem todas as variáveis precisam ter a mesma ordem de integração para existir cointegração. Supondo o caso de três variáveis, $x_t = [y_t, x_t, z_t]'$, sendo duas integradas de ordem 2, y_t e x_t, e uma integrada de ordem 1, z_t, é possível imaginar que as duas variáveis de maior ordem sejam C(2,1) e a resultante, quando combinada a z_t seja C(1,1). Isto é, pode-se imaginar um vetor β entre y_t e x_t que gere uma variável, w_t I(1), e, a seguir, um vetor π que combine w_t e z_t, resultando em um modelo estacionário. Formalmente, as contas são: [Bueno, 2011]

$$\frac{1}{\pi_1 y_t} + \frac{\beta_0}{\pi_1} x_t = w_t \quad \sim I(1)$$

$$\pi_1 w_t + \pi_2 z_t = \mu_t \quad \sim I(0).$$

A conclusão desta abstração é a necessidade de haver pelo menos duas variáveis integradas de mesma ordem na ordem máxima de integração entre todas as variáveis para que exista cointegração.

4.7 Teste para Cointegração

O teste de cointegração utilizado neste trabalho é o de Engle-Granger que conforme [Bueno, 2011] é um teste imediato e de fácil implementação. Suponha um sistema bivariado nas variáveis Y_t e X_t, integradas de ordem 1. A metodologia consiste em estimar a relação de longo prazo e

armazenar os resíduos. Se as variáveis forem cointegradas, os resíduos serão estacionários. Para tal, o teste é realizado nas seguintes etapas:

1. Executar o teste de raiz unitária nas variáveis de interesse e certificar que são I(1);

2. Estimar a relação de longo prazo e obter $\hat{\mu}_t$;

3. Realizar o teste de raiz unitária nos resíduos estimados, usando o procedimento ADF:

$$\Delta\hat{\mu}_t = \alpha\hat{\mu}_{t-1} + \sum_{i=1}^{p-1} \lambda_{i+1}\Delta\hat{\mu}_{t-i} + \nu_t$$

A não rejeição de $H_0 : \alpha = 0$ implica que os resíduos têm raiz unitária, de modo que as variáveis não cointegram.

5 Aplicação Empírica

5.1 Introdução

Neste última etapa do trabalho, será apresentado o passo a passo para a construção de uma estratégia de investimentos em criptomoedas. Como o mercado de criptomoedas é prematuro e ainda possui diversas ineficiências, optou-se pela construção de uma técnica de arbitragem estatística. Além disso, por acreditar-se no potencial de longo prazo das criptomoedas, dadas as tecnologias e revoluções que o *Blockchain* traz para a sociedade, a estratégia será focada em aumentar a quantidade de Bitcoins e não necessariamente no aumento de capital a curto prazo.

5.2 Coleta e Tratamento dos Dados

Dentro do universo de corretoras e criptomoedas, optou-se por trabalhar com a corretora Poloniex. Vale observar que, como o mercado de criptomoedas não é regulado e não tem uma entidade central que controle os preços, há divergências e discrepâncias de preços entre diferentes corretoras. Além disso, como pretende-se trabalhar com uma estratégia de arbitragem estatística, na qual são abertas posições simultâneas de compra (*long*) e venda (*short*), é necessário que os ativos permitam operações alavancadas. Com isso esclarecido, selecionaram-se 10 criptomoedas listadas na Poloniex que atendem os seguintes requisitos:

- (a) Possuir Alta Liquidez[1];
- (b) Ser negociada diretamente contra o Bitcoin;
- (c) Permitir a alavancagem[2] das operações.

Após esta definição inicial, para a obtenção dos dados históricos, foi desenvolvida uma conexão API (*Application Programming Interface*)

[1] Liquidez é a facilidade de converter o ativo em dinheiro.

[2] Alavancagem é quando não se utiliza o capital próprio para operar no mercado, mas toma-se um empréstimo.

diretamente com a Poloniex, uma vez que, como comentando, o preço oscila entre corretoras e deseja-se obter o resultado mais fidedigno possível. Então, para desenvolver esta conexão utilizou-se programação PHP e *JavaScript* e, na sequência, executou-se o programa através do Servidor XAMPP. A seguir, na figura 5.1, segue um exemplo de código utilizado para realizar o procedimento:

```
public function returnChartData($currencyPair, $start, $end, $period)
{
        if ($currencyPair === null) {
        $currencyPair = 'all';
        }

        return $this->queryPublic('returnChartData&currencyPair='.$currencyPair.'&start='.$start.'&end='.$end.'&period='.$period);
}
```

Figura 5.1: Exemplo código API
Fonte: Autor

Por fim, obtiveram-se os dados dos ativos para o período de 5 de outubro de 2015 e 13 de setembro de 2017. Os dados foram coletados em períodos de 4 em 4 horas, ressaltando que o mercado de criptomoedas não dorme, estando ativo 24 horas por dia, 7 dias por semana.

A seguir, segue a lista de criptomoedas selecionadas.

- BTS BitShares;
- DASH Dash;
- DOGE Doge;
- ETH Ethereum;
- FCT Factom;
- LTC LiteCoin;
- MAID MAID;
- STR Stellar Lumens;
- XMR Monero;
- XRP Ripple;

5.3 Análise Inicial

Posteriormente à coleta dos dados, analisou-se o comportamento das 10 criptomoedas selecionadas, para entender os padrões e relações das mesmas. Assim, torna-se mais concreta a construção da estratégia. Para tal, observou-se conjuntos de moedas em torno da mesma faixa de preço.

No primeiro grupo têm-se moedas com custo entre 20 e 250 microbitcoins[3]. Percebe-se que a moeda $MAID$ apresenta um padrão que aparentemente não apresenta relação com XRP e BTS. Além disso, cabe destacar que esta criptomoeda tem apresentado comportamento,

[3] μ = 1 microbitcoin = 0.00000010 BTC

ao que tudo indica, estacionário. Portanto, a melhor estratégia para a mesma é desenvolver uma técnica de retorno à média diretamente sobre o seu preço.

Ao analisar as moedas *Ripple*(XRP) e *BitShares*(BTS), percebeu-se que, aparentemente, as duas apresentam co-movimentos e tendem a manter uma relação de estabilidade entre seus preços, ver figura 5.2. Logo, este par de moedas é um provável candidato a apresentar cointegração. Ainda assim, deve-se observar que o movimento do *BitShares* segue o *Ripple* com uma defasagem de tempo, o que também indica que, talvez, ao utilizar um modelo preditivo do preço do BTS em função do XRP, bons retornos poderiam ser obtidos.

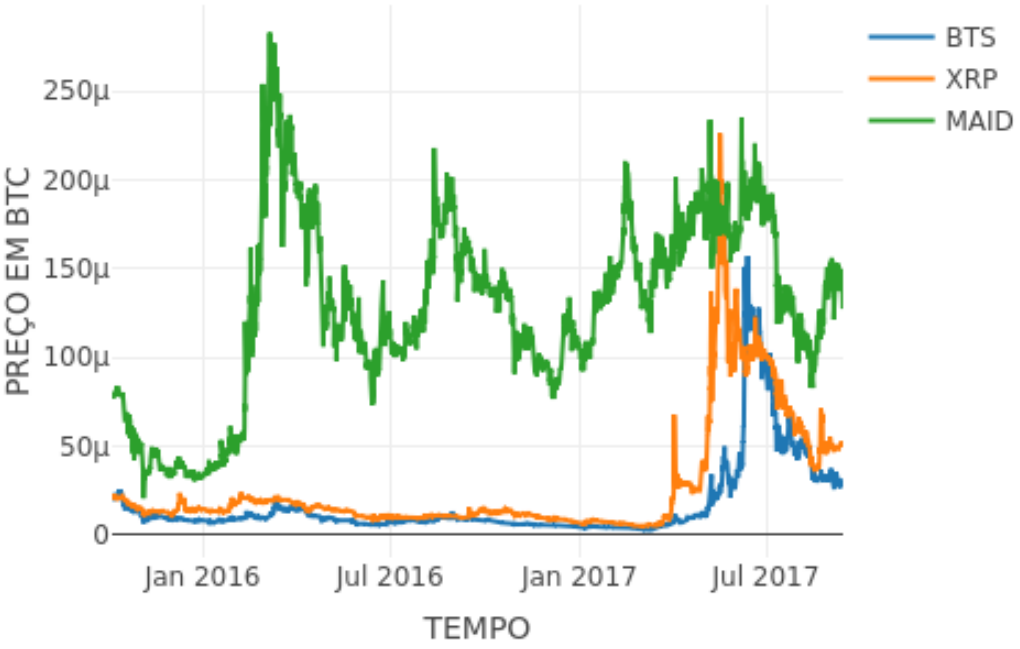


Figura 5.2: Variação do preço BTS, XRP e MAID
Fonte: Autor

Tanto o *Dash*, com sua filosofia de anonimato, velocidade e segurança, quanto o *Ethereum*, com seus contratos inteligentes, são dois gigantes do mundo das criptomoedas e estão entre os maiores em capitalização de mercado. Ao observar os seus preços históricos no gráfico 5.3, constata-se que as duas moedas aparentemente estão em uma "briga", na qual, quando uma delas sobe, a outra não só sobe, mas também esforça-se para ultrapassar os preços da concorrente. Essa ligação entre estas criptomoedas é bastante interessante, pois tem-se um par que presumivelmente não seja correlacionado, mas que deve apresentar uma estabilidade entre seus preços a longo prazo, dado que ambas buscarão estar sempre no mesmo patamar de preço.

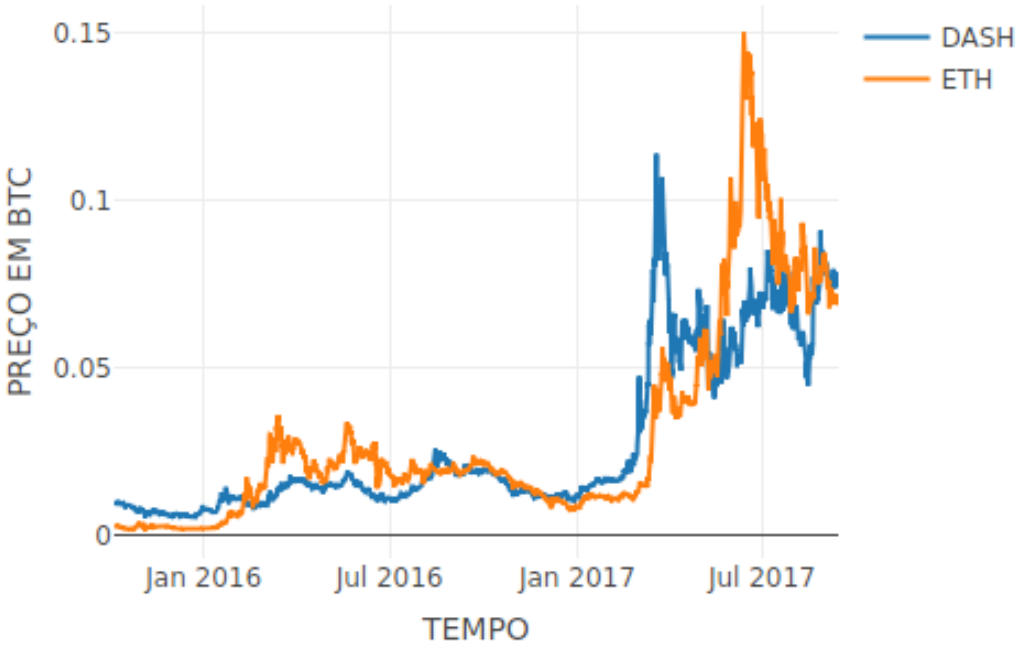


Figura 5.3: Variação do preço DASH e ETH
Fonte: Autor

Ao examinar o *Monero*, *Litecoin* e *Factom* não encontrou-se nenhum sinal aparente que estas criptomoedas pudessem apresentar uma relação entre si, conforme figura 5.4. Isso se justifica pelo fato de que as tecnologias e focos destas criptomoedas são bem diferentes: enquanto o *Monero* tem foco na privacidade; o *Factom* almeja o desenvolvimento de uma plataforma para preservar, garantir e validar ativos digitais; já o *Litecoin* procura melhorar alguns dos problemas do *Bitcoin*, como a velocidade de transação e a mineração de moedas.

Figura 5.4: Variação do preço XMR, LTC E FCT
Fonte: Autor

Ao analisar a figura 5.5: *Doge* , *Stellar* e *Lumens* não apresentam nenhuma relação independentemente da análise que seja realizada. Contudo, vale destacar que, apesar de estar em faixas de preços totalmente diferentes, o STR apresenta comportamento semelhante ao BTS e XRP vistos anteriormente. Isto é totalmente plausível, dado que estes 3 projetos são focados em prover serviços bancários.

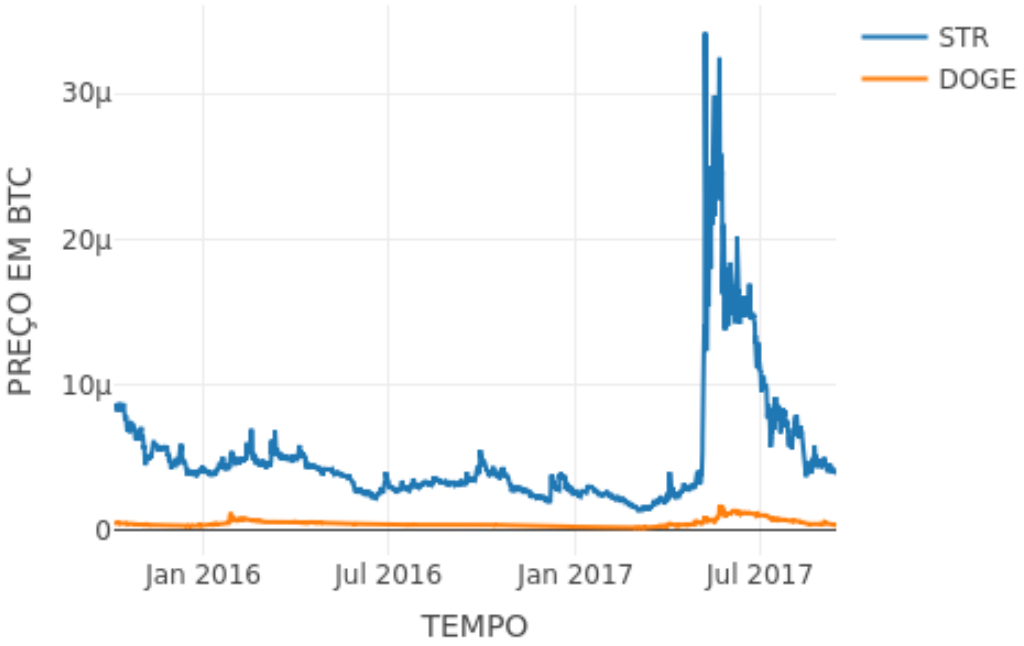
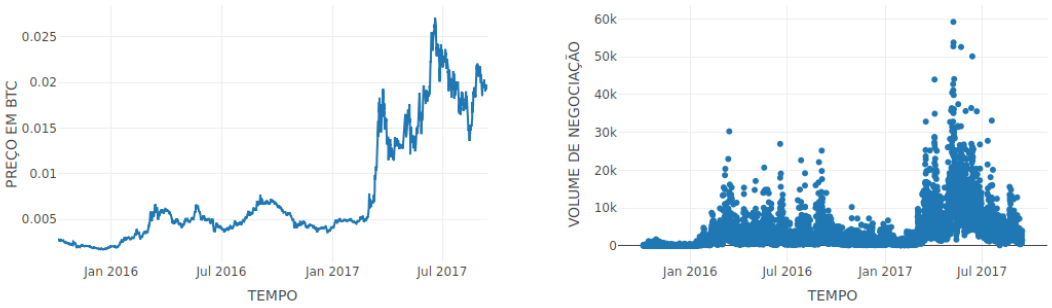


Figura 5.5: Variação do preço STR e DOGE
Fonte: Autor

Por fim, criou-se uma variável que representa o mercado destas 10 criptomoedas como um todo. Para tal, a média dos preços destas moedas foi feita, ponderada pelo volume de transação de cada uma. Ao analisar o histórico de preços desses ativos, percebe-se que há um movimento totalmente atípico, aproximadamente entre abril e agosto de 2017, no qual há um crescimento totalmente descontrolado do mercado, assim como extrema volatilidade[4] (que em criptomoedas já não é baixa em outros momentos), de acordo com a figura 5.6. Este fenômeno também pode ser observado no volume de negociações realizadas. Além disso, percebe-se que ao final de setembro, o padrão de negociações está voltando aos níveis anteriores a abril de 2017. Como a técnica de arbitragem proposta por este trabalho usa alavancagem, sendo extremamente sensível a grandes oscilações de preço, a abertura de posições só será realizada em momentos em que o volume encontra-se estável. Ao sair desta estabilidade, não se abrirão posições no mercado.

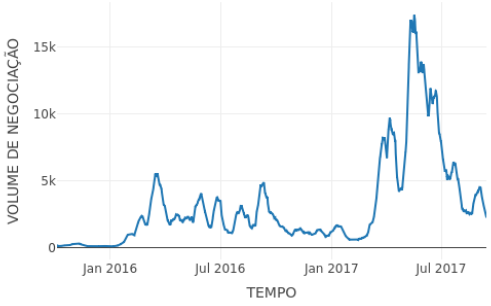
Por este viés, na próxima seção será analisado se há coerência em realizar esta divisão da estratégia através de testes para verificar se existe uma quebra estrutural nos preços e no volume de negociações.

[4]Volatilidade é uma medida de dispersão dos retornos de um determinado ativo.



(a) Preço Mercado

(b) Volume Mercado



(c) Volume Médio Mercado

Figura 5.6: Resumo do preço e volume médio do Mercado
Fonte: Autor

5.4 Pontos de Quebra Estrutural

O teste para estimar o número de quebras estruturais existentes em uma série temporal, desenvolvido por Achim Zeileis e disponibilizado no pacote "*strucchange*"do *software* R, foi utilizado para determinar possíveis pontos em que a série do preço geral das criptomoedas apresente possíveis quebras estruturais. Este teste verifica computacionalmente a estabilidade da série ao incluir um número "n"de quebras, assim o número ótimo de quebras é baseado em uma abordagem de programação dinâmica. Por fim, são apresentados os valores do Critério de Informação Bayesiano (BIC) e da Soma dos Quadrados dos Resíduos (RSS), onde quanto menor estes valores forem mais estabilidade a série terá. Com isso explicado, percebe-se, na imagem 5.7, que ao incluir-se apenas uma quebra da série, o BIC e RSS caem violentamente. Ademais, a partir desta quebra, independentemente de quantas outras forem adicionadas, o BIC e o RSS se mantém praticamente constantes, o que indica que ao incluir somente uma quebra estrutural a série torna-se estável. Pela tabela 5.1 apresentada, verifica-se que esta quebra acontece em torno do ponto 6322, referente a 13 março de 2017, o que é coerente com a discussão apresentada anteriormente.

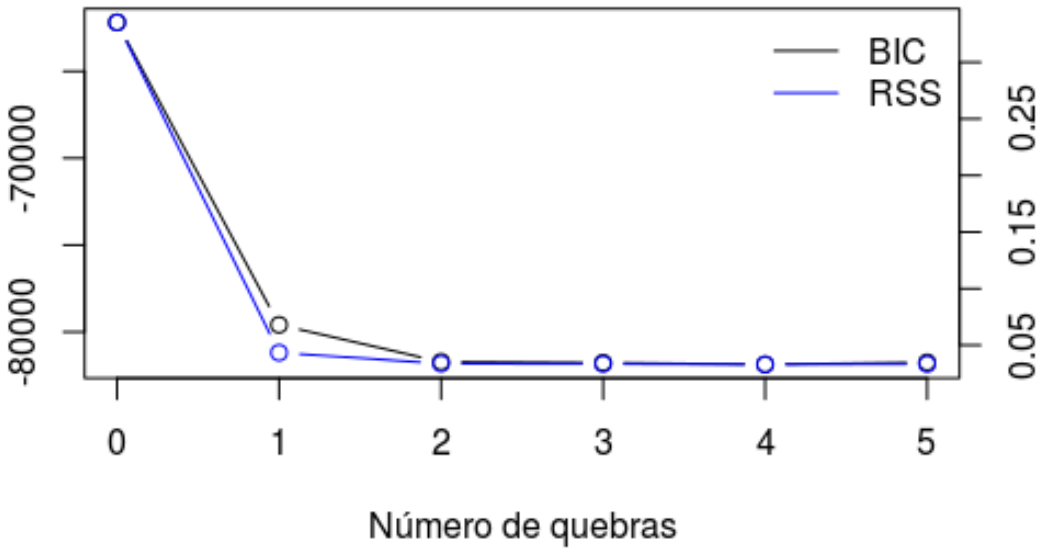


Figura 5.7: Número de quebras na série do preço
Fonte: Autor

Tabela 5.1: Pontos de quebra na série do preço

Nº Quebras	Pontos de Quebra				
1				6322	
2	1784			6329	
3	1792		4639	6323	
4	1784	3612	4889	6323	
5	1784	3408	4685	5962	7239

Fonte: Autor

Fora o teste apresentado, também verificou-se possíveis pontos de quebra na série do volume de negociações. Ao analisar este segundo resultado, conforme a figura 5.8, percebe-se que a interpretação é análoga ao teste com o preço, o que confirma a necessidade de haver diferentes tratamentos antes e depois deste período. Um fator importante a se atentar é que os pontos de quebras definidos a partir do volume acontecem sempre alguns passos à frente daqueles que são definidos pelo preço, conforme a tabela 5.2. Com essa informação pode-se pensar no desenvolvimento de um modelo para a previsão de quebras estruturais a partir do volume de negociação.

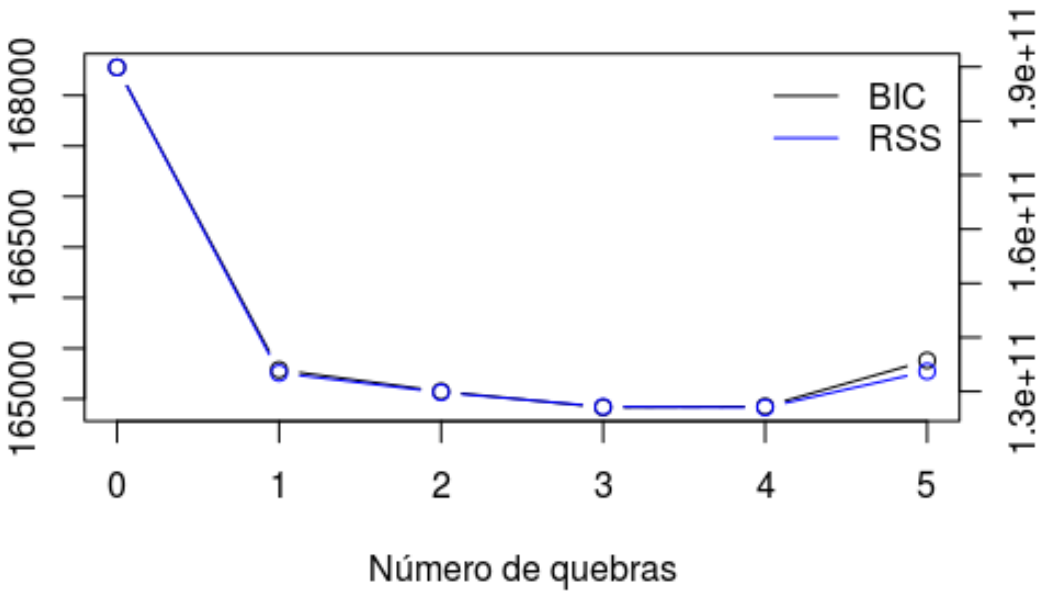


Figura 5.8: Número de quebras na série do volume
Fonte: Autor

Tabela 5.2: Pontos de quebra na série do volume

Nº Quebras	Pontos de Quebra				
1				6274	
2	1507			6274	
3	1527		4248	6274	
4	1527	2847	4248	6274	
5	1527	2847	4248	5962	7239

Fonte: Autor

5.5 Teste de Estacionariedade

Nesta etapa, todas as séries já estavam em conformidade com a quebra estrutural definida anteriormente. Todavia, ainda era necessário verificar quais séries de criptomoedas poderiam ser cointegradas, ou seja era preciso aferir quais delas eram não estacionárias. Para tal, realizou-se o teste de Dickey-Fuller Aumentado (ADF), considerando a quebra estrutural dos processos, assim definindo quais séries são estacionárias ou não. Os resultados são apresentados na tabela 5.1. Além disso, vale mencionar que as séries não estacionárias, por serem referentes a ativos financeiros, são I(1), conforme definição de [Bueno, 2011]. Logo, todas têm a mesma ordem de integração.

Tabela 5.3: p-valores do teste ADF das séries de preços das criptomoedas

Ativo	p-valor	Decisão
BTS	0.01447	Estacionária
DASH	0.99	Não Estacionária
DOGE	0.1072	Não Estacionária
ETH	0.6382	Não Estacionária
FCT	0.3484	Não Estacionária
LTC	0.0441	Estacionária
MAID	0.6445	Não Estacionária
STR	0.01	Estacionária
XMR	0.4361	Não Estacionária
XRP	0.149	Não Estacionária

Fonte: Autor

5.6 Procurando a Cointegração

Após definir-se quais séries são estacionárias, ou não, testaram-se combinações das mesmas para verificar quais apresentavam a propriedade da cointegração. Nesse processo, também incluíram-se combinações com séries estacionárias. Das combinações testadas, a que apresentou melhores resultados inclui as seguintes moedas: *Ethereum*, *Ripple*, *BitShares* e *Stellar*. Esta combinação não foi eleita somente por seus princípios matemáticos (os resíduos da regressão são estacionários, teste da raiz unitária possui um p-valor < 0.01), mas também pelos fundamentos atrás destas moedas, dado que *Ethereum* e *Ripple*, ambos I(1), são concorrentes de longa data para obterem o segundo lugar no pódio de capitalização de mercado, além disso as moedas *BitShares* e *Ripple*, que são I(0), foram incluídas pois atuam em um mesmo segmento que o *Ripple* e, conforme mostrado anteriormente na figura 5.2, apresentam movimentos semelhantes em seus preços.

Como os resíduos produzidos pela estimativa do modelo MQO são estacionários, os mesmos devem sempre retornar à média[5], dando margem para o desenvolvimento de uma técnica que possibilite obter lucros do mercado, pois há uma previsibilidade e um padrão do que deve acontecer. o resíduo torna-se um ativo sintético que é dado pela combinação das 4 moedas apresentadas. Segue a descrição desta relação.

$$AtivoSintetico_t = \mu_t = \beta_1 ETH_t + \beta_2 XRP_t + \beta_3 BTS_t + \beta_4 STR_t$$

Um ponto importante a se destacar aqui é que o modelo construído foi eleito ao testar diversas combinações de séries e alterando-se a quanti-

[5] A reversão média é uma teoria que sugere que os preços e os retornos de um ativo eventualmente se movem para a média.

dade de ativos na composição. Contudo, existem alternativas mais inteligentes para selecionar criptomoedas que apresentem a propriedade da cointegração como, por exemplo, a aplicação de um Modelo Vetorial de Correções de Erros (MVCE).

Ao averiguar a auto-correlação do ativo sintético na figura 5.9, percebe-se que o mesmo apresenta dependência no tempo. Então, algumas possibilidades foram exploradas em busca de um melhor ajuste do modelo.

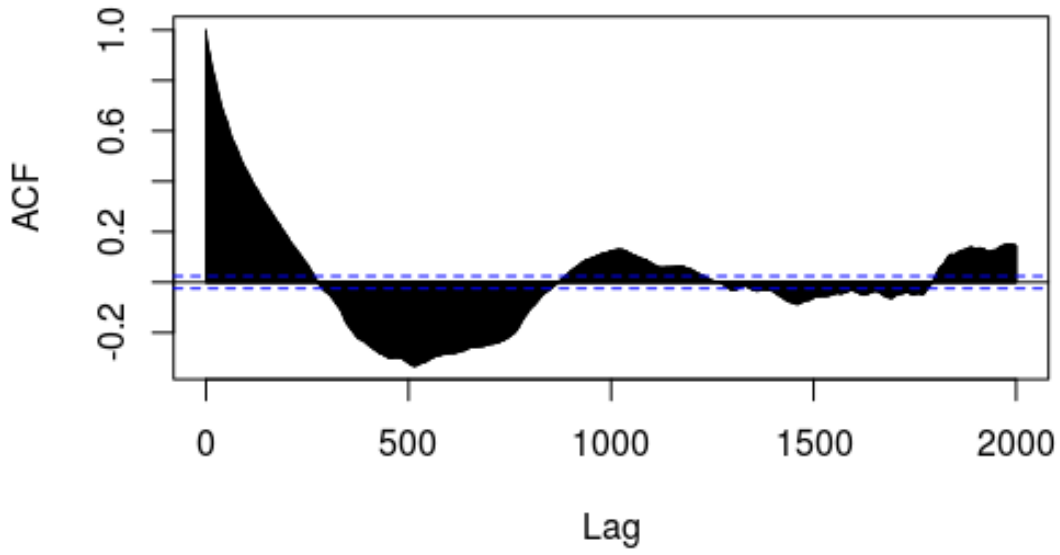


Figura 5.9: Autocorrelação do Ativo Sintético
Fonte: Autor

Com isso, verificou-se que o ativo sintético se trata de um auto-regressivo de ordem 4 e, portanto, o preço em um instante t é dependente do preço nos últimos 4 instantes de tempo. Isso pode ser justificado pois o preço atual é altamente influenciado pelo preço no instante anterior, e a medida que o ativo sintético se desenvolve no tempo, esta influência não faz mais efeito.

Criou-se um novo modelo considerando os quatro preços anteriores do ativo sintético e, assim, constatou-se pelo teste de $Box - Pierce$ que os novos resíduos são IID (p-valor = 0.7121), o que também pode ser percebido nos gráficos de autocorrelação na sequência (figura 5.10).

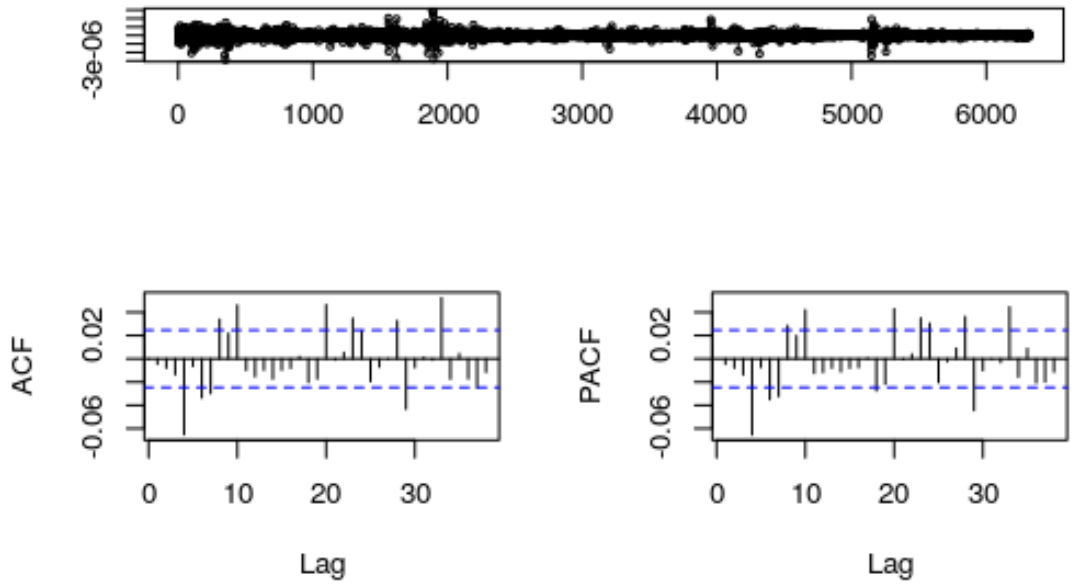


Figura 5.10: Resíduos dos termos Auto-Regressivos do Ativo Sintético
Fonte: Autor

Por fim, a figura 5.11 mostra a série histórica dos preços do ativo sintético construído, a partir das 4 moedas escolhidas.

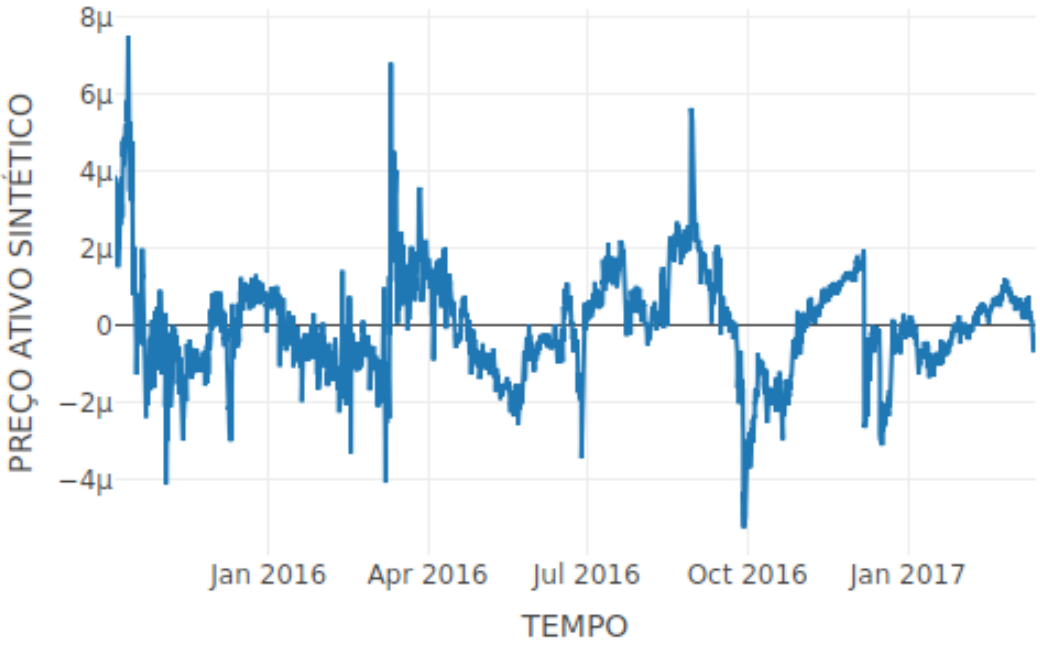


Figura 5.11: Preço do Ativo Sintético
Fonte: Autor

5.7 Construção da Estratégia

Nesta etapa, como já foram definidas as criptomoedas que compõem a técnica de arbitragem estatística, assim como o peso (β da equação) que elas teriam na composição do ativo sintético, serão determinados os pontos de entrada e saída da estratégia. No intuito de atingir este objetivo, padronizou-se o ativo sintético e definiu-se que quando o mesmo apresentar 2 desvios, tanto para cima quanto para baixo da média, deve-se abrir uma posição. Nos casos onde o ativo sintético estiver 2 desvios abaixo da média, as moedas com βs positivos devem ser compradas e devem ser vendidas aquelas com βs negativos. Em casos opostos (2 desvios acima da média), deve-se realizar a operação contrária. Posteriormente, deve-se manter esta posição aberta até que o ativo sintético retorne para a média, ou seja, no ponto 0, para liquidar a posição e assim obter os retornos. Uma visualização da estratégia pode ser encontrada na figura 5.12.

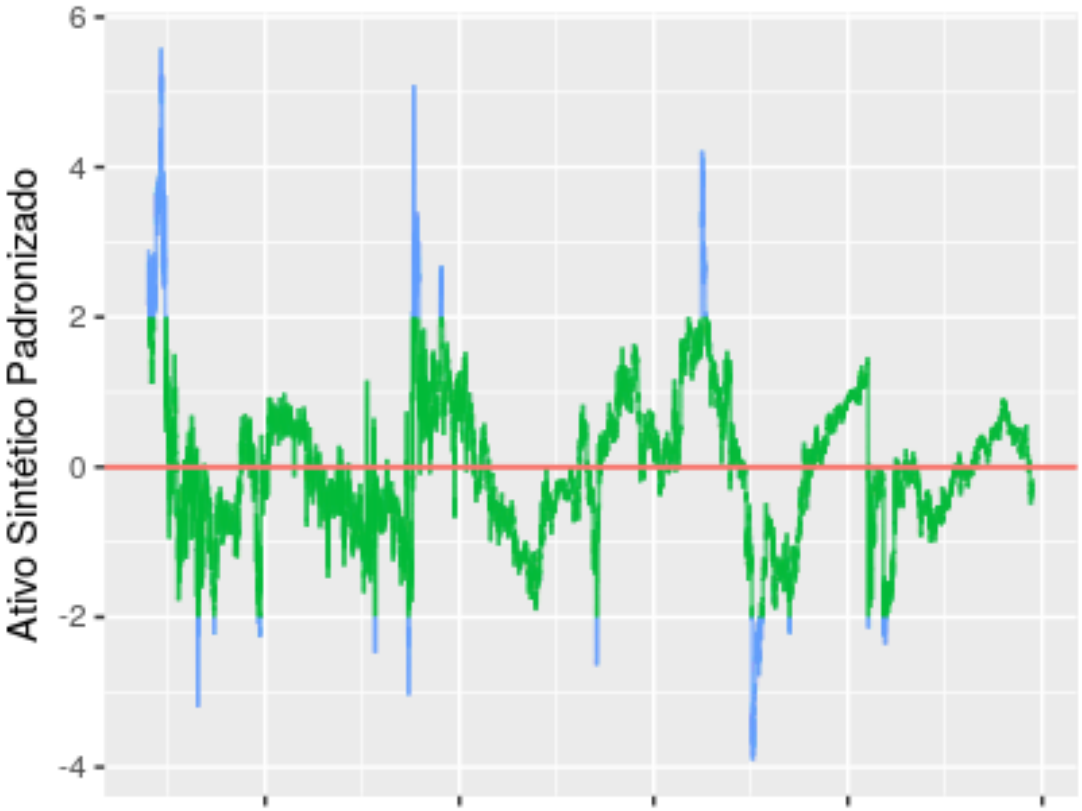


Figura 5.12: Pontos de entrada e saída
Fonte: Autor

5.8 Performance da Estratégia

Ao final de tudo, procurou-se apurar qual a performance da estratégia construída. Portanto, criou-se uma série histórica dos retornos acumulados, podendo ser descritos pela fórmula a seguir:

$$\text{Retorno}_t = \begin{cases} Retorno_{t-1}, & \text{nao posicionado} \\ Retorno_{t-1}(1 + \frac{\sum_{i=1}^{4} \beta_i (Preco_{t,i} - Preco_{t-1,i}) Indicadora_{t,i}}{\sum_{i=1}^{4} |\beta_i| Preco_{t-1,i}}, & \text{caso contrário.} \end{cases}$$

Nessa equação, a variável Preço contém a cotação das 4 criptomoedas selecionadas para o período analisado e a variável Indicadora serve para definir se a moeda está comprada ou vendida em determinado instante. Quando a moeda está comprada a variável Indicadora assume valor 1, caso contrário assume valor -1.

Abaixo tem-se o gráfico da série dos retornos, obtendo aproximadamente 80% de ganhos no período entre 5 de outubro de 2015 e 13 de março de 2017, o que indica um bom desempenho da estratégia desenvolvida. Vale lembrar que este retorno é sobre o *Bitcoin*, e que o mesmo apresentou uma valorização entre 5 de outubro de 2015 e 13 de setembro de 2017, aproximada de 1600%. Assim, o ganho real da estratégia supera os 3000%. Contudo, apesar do excelente retorno, O *Sharpe Ratio* da estratégia é baixo (apenas 0.01528346), o que explica-se pela extrema volatilidade do mercado de criptomoedas. Com isso, deve-se ter muita prudência e cuidados com a alavancagem utilizada ao adotar esta estratégia ou qualquer outra neste mercado, pois os riscos de perdas são altíssimos.

Na figura 5.13, também é possível visualizar períodos em que a estratégia está com posições ativas (1) ou não (0). É importante perceber que a mesma tem períodos longos, às vezes chegando a meses, em que não existem oportunidades de arbitragem. Sabendo disso, vale frisar que a paciência deve andar lado a lado com todo o investidor.

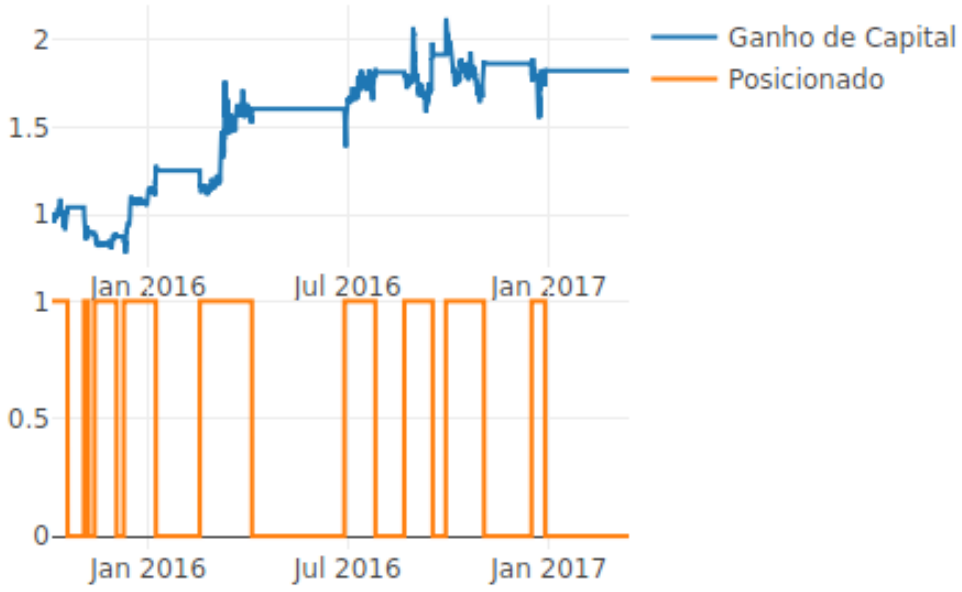


Figura 5.13: Ganho de Capital
Fonte: Autor

6 Considerações finais

Ao final deste trabalho, acredita-se que o mesmo conseguiu apresentar de maneira satisfatória o que são as criptomoedas, assim como sua importância e as mudanças de paradigmas e tecnologias que as mesmas trazem para sociedade. Além disso, foram explorados conceitos e técnicas de arbitragem estatística, o que acaba por expandir a compreensão e o conhecimento sobre os mercados financeiros. Fora isto, o trabalho mostrou que técnicas de arbitragem estatística apresentam bons resultados no mercado de criptomoedas.

Contudo, há ressalvas em relação à validade dos resultados apresentados, pois o período analisado ainda é curto, e o próprio mercado de criptomoedas ainda é muito novo. Isso indica que a cointegração entre estes ativos pode ser espúria e não se refletir em uma relação de estabilidade a longo prazo no preço destes ativos. Portanto, caso a técnica de arbitragem continue sendo utilizada, é necessário revisar periodicamente se as propriedades de cointegração se mantêm e, em caso negativo, deve-se liquidar as posições.

Também se faz necessária a previsão de possíveis quebras estruturais futuras nessa série, como já comentado, uma opção seria utilzar um modelo GARCH para fazer esta previsão a partir do volume de negociações. Além disso, é interessante que se entenda o comportamento do mercado nos períodos em que houver a quebra estrutural, uma ideia seria posicionar todos os ativos para a mesma direção que o mercado está se movendo, e, com isso, melhorar ainda mais os resultados.

Por fim, é necessário dizer que independentemente do preço e dos retornos financeiros obtidos com criptomoedas, as tecnologias que as mesmas trazem terão grande impacto no futuro e irão remodelar a economia global. A primeira revolução digital trouxe a internet da informação. A segunda, por meio do *Blockchain*, está trazendo Internet do valor: uma nova plataforma descentralizada que pode nos ajudar a tornar o mundo em um lugar melhor. Talvez, daqui a 10, 20, ou 30 anos tanto o *Bitcoin*, quanto outras criptomoedas apresentadas no trabalho não existam mais e sejam apenas uma história, mas certamente o *Blockchain* estará mais vivo do que nunca, dado a quantidade massiva de dados produzidos segundo a segundo, o desenvolvimento da internet das coisas, a crescente demanda pela economia compartilhada, a busca da descentralização do poder, e de outros motivos. Assim, este trabalho encerra frisando a importância da comunidade científica se adaptar e estudar a revolução que estamos vivendo, para que possa participar ativamente da mesma e não apenas permanecer como espectadora.

Referências Bibliográficas

Bitcoin: A peer-to-peer eletronic cash system. *Bitcoin Org*, 2008.

Andreas M. Antonopoulos. *Mastering Bitcoin: Unlocking Digital Crypto-Currencies*. O'REILLY, 2014.

Anton Badev. Bitcoin: Technical background and data analysis. *Finance and Economics Discussion Series Divisions of Research Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.*, 2014.

Rodrigo de Losso da Silveira Bueno. *Econometria de Séries Temporais*. Cengage Learning, 2011.

João F. Caldeira. Arbitragem estatística, estratégia long-short pairs trading, abordagem com cointegração aplicada ao mercado de ações brasileiro. 2013.

David A. Dickey. Distribution of estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 1979.

Robert F. Engle. Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 1987.

Ryan Farell. An analysis of the cryptocurrency industry. *Wharton Research Scholars*, 2015.

Ruben Joakin Gundersen. Statistical arbitrage: High frequency pairs trading. Master's thesis, Norwegian School of Economics, 2014.

Tarcísio Saraiva Rabelo Junior. Mercados eficientes e arbitragem: um estudo sob o enfoque das finanças comportamentais. *Revista Contabilidade Finanças*, 2004.

Risk Latte. Morgan stanley and the birth of statistical arbitrage. 2011.

Arvind Narayanan. *Bitcoin and Cryptocurrency Technologies*. Princeton University Press, 2016.

Luiz Paulo Rodrigo de Freitas Parreira. Arbitragem estatística e inteligência artificial. Master's thesis, Universidade Federal de São Paulo, 2007.

Thaís Neves Pereira. Cointegração: Uma relação de equilíbrio de longo prazo. Master's thesis, Universidade Federal do Rio Grande do Sul, 2013.

Doles Silva. Cryptocurrencies: International regulation and uniformization of practices.

Isak Starlander. Counterparty credit risk on the blockchain. Master's thesis, KTH Institute of Technology School of Engineering Sciences, 2017.

Florian. Tschorsch. Bitcoin and beyond: A technical survey on decentralized digital currencies. *Humboldt University of Berlim*, 2015.

Ganapathy Vidyamurthy. *Pairs Trading - Quantitative Methods and Analysis*. Wiley Sons, 2004.