

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
DEPARTAMENTO ACADÊMICO DE MECÂNICA  
CURSO DE ENGENHARIA MECÂNICA

GUSTAVO DOTTI GRANDO  
VINICIUS NUNES SILVA RIBEIRO

**PROPOSTA DE UM PROCEDIMENTO PARA ANÁLISE DE BIG DATA  
POR MEIO DE MÉTODOS ESTATÍSTICOS.**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA

2019

GUSTAVO DOTTI GRANDO  
VINICIUS NUNES SILVA RIBEIRO

**PROPOSTA DE UM PROCEDIMENTO PARA ANÁLISE DE BIG DATA  
POR MEIO DE MÉTODOS ESTATÍSTICOS.**

Monografia do Projeto de pesquisa apresentada à disciplina de Trabalho de Conclusão de Curso - Tcc2 do Curso de Engenharia Mecânica da Universidade Tecnológica Federal do Paraná, como requisito parcial para aprovação na disciplina.

Orientadora: Profa. Dra. Cleina Yayoe Okoshi

CURITIBA

2019

## TERMO DE APROVAÇÃO

Por meio deste termo, aprovamos a monografia do Projeto de Pesquisa "PROPOSTA DE UM PROCEDIMENTO PARA ANÁLISE DE BIG DATA POR MEIO DE MÉTODOS ESTATÍSTICOS", realizado pelo aluno(s) Gustavo Dotti Grando e Vinicius Nunes Silva Ribeiro, como requisito para aprovação na disciplina de Trabalho de Conclusão de Curso 2, do curso de Engenharia Mecânica da Universidade Tecnológica Federal do Paraná.

Prof. Dra Cleina Yayoe Okoshi

DAMEC, UTFPR

Orientador

Prof. Dr Leandro Magatão

CPGEI, UTFPR

Avaliador

Prof. M.e Walter Luís Mikos

DAMEC, UTFPR

Avaliador

Curitiba, 28 de Novembro de 2019.

## RESUMO

Grando, Gustavo Dotti; Ribeiro, Vinicius Nunes Silva. Proposta de um procedimento para análise de Big Data por meio de métodos estatísticos. 51 f. Trabalho de conclusão de curso – Tcc2, Bacharelado em Engenharia Mecânica, Departamento Acadêmico de Mecânica, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

A Indústria 4.0 também chamada de quarta revolução industrial, é um conceito moderno de fábricas inteligentes que possui sua base em diversas tecnologias modernas e emergentes como o Analytics e Big data, esta tecnologia sendo transformada pelo blockchain, inovando em descentralizar, criptografar e tornando público o armazenamento de informação. O presente trabalho tem como objetivo propor um procedimento de análise de big data utilizando métodos estatísticos e aplicando a uma base de dados de criptomoedas, sendo essa uma big data acessível a todos e com sua principal característica ter origem na tecnologia do blockchain. Será feita revisão bibliográfica dos métodos estatísticos, a escolha da base de dados, preparação da base suas variáveis de análise, aplicação do procedimento através do software de análise e tendo como resultado um procedimento claro de análise de dados de big data utilizando estatística e a análise da base de dados de 2 anos das principais criptomoedas atualmente do mercado. O principal resultado foi a análise estatística das principais criptomoedas, onde identificou a correlação entre elas e também a influência da principal criptomoedas em relação as outras moedas. Também destacou que o procedimento proposto é aplicável em big data.

**Palavras-chave:** Análise em Big Data. Análise estatística. Indústria 4.0. Blockchain. Criptomoedas.

## ABSTRACT

Grando, Gustavo Dotti; Ribeiro, Vinicius Nunes Silva. Proposal of a procedure for Big Data analysis using statistical methods. 51 p. Undergraduate Thesis, Mechanical Engineering, Academic Department of Mechanical, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

The Industry 4.0 also called the fourth industrial Revolution, it's a modern concept of smart factories that has in your foundation many modern and emergent technologies as the Analytics and Big data, a technology that has being transformed by the blockchain technology. Innovating in decentralize, cryptograph and making it public, it is transforming the way of storing data. This paper has the objective of propose a procedure of analysis of the big data using statistics methods and applying it to a cryptocurrency data set, a big data set that is available to anyone and have your foundation at the blockchain. It will be done the literature review of the statistics methods, the choice of data set, preparation of the used variables in the data set and applying it at the analytics software having for results an procedure of analysis of big data dataset using the statistics and an analysis of two years dataset of the mainly cryptocurrencies in the market. The mainly result was the statistical analysis of the mainly cryptocurrency in relation to the other coins. Also it highlighted the proposed procedure is applicable to big data datasets.

**Keywords:** Big Data analytics. Statistic analysis. Industry 4.0. Blockchain. Cryptocurrencies.

## LISTA DE FIGURAS

Figura 1 – Indústria 4.0 e suas tecnologias .....	15
Figura 2 – Os 5V's do Big Data .....	17
Figura 3 - Esquema de rede ponto-a-ponto (peer-to-peer) .....	18
Figura 4 - Estrutura do Blockchain .....	19
Figura 5 – Relação linear entre X e Y .....	24
Figura 6 – Gráfico da equação modelo .....	25
Figura 7 – Fluxograma da metodologia .....	28
Figura 8 – Procedimento de análise de Big Data para Blockchain .....	31
Figura 9 – Fonte dos dados coin market cap .....	34
Figura 10 – Identificação de Outliers .....	37
Figura 11 – Tabelas de resultados de análise de variância .....	42
Figura 12 – Gráfico de linha ajustado da moeda Ethereum .....	42
Figura 13 – Comparação Real e Previsão .....	44

## LISTA DE TABELAS

Tabela 1 – Tabela de interpretação da correlação .....	24
Tabela 2 – Ranking das criptomoedas. Valores referentes a 21/09/2019 .....	33
Tabela 3 – Nova coluna com a variação .....	35
Tabela 4 – Nova planilha.....	36
Tabela 5 – Completando os dados faltantes .....	36
Tabela 6 – Correlação sem retirar os Outliers.....	38
Tabela 7 – Correlação após a retirada dos Outliers .....	38
Tabela 8 – Resultados da análise descritiva .....	39
Tabela 9 – Resultado variância Ethereum Bitcoin.....	40
Tabela 10 – Resultados de regressão.....	43
Tabela 11 – Resultado de previsão de movimento de mercado.....	44

## **LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS**

BTC – Bitcoins.

CPS – Cyber-Physical Systems.

IA – Inteligência Artificial.

IDC – International Data Corporation.

IoT – Internet of Things.

# SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	11
1.1. OBJETIVOS .....	12
1.1.1. Objetivo geral.....	12
1.1.2. Objetivos específicos.....	12
1.2. JUSTIFICATIVA .....	13
<b>2. FUNDAMENTAÇÃO TEÓRICA</b> .....	14
2.1. INDÚSTRIA 4.0.....	14
2.2. BIG DATA .....	15
2.3. BLOCKCHAIN.....	17
2.4. CRIPTOMOEDAS .....	19
2.5. ESTATÍSTICA DESCRITIVA.....	21
2.5.1. Média aritmética .....	21
2.5.2. Mediana.....	22
2.5.3. Variância.....	22
2.5.4. Desvio padrão.....	23
2.6. ANÁLISE DE CORRELAÇÃO .....	23
2.7. ANÁLISE DE REGRESSÃO LINEAR.....	24
2.7.1. Método dos mínimos quadrados.....	26
2.7.2. Coeficiente de determinação $R^2$ .....	26
<b>3. METODOLOGIA</b> .....	28
3.1. REVISÃO BIBLIOGRÁFICA DOS CONTEÚDOS .....	28
3.2. IDENTIFICAÇÃO DAS FERRAMENTAS ESTATÍSTICAS.....	28
3.3. IDENTIFICAÇÃO DA BASE DE DADOS DE BIG DATA ACESSÍVEL E APROPRIADA PARA O ESTUDO (BLOCKCHAIN – CRIPTOMOEDAS) .....	29
3.4. TREINAMENTO DE UTILIZAÇÃO DOS SOFTWARES.....	29
3.5. DESENVOLVIMENTO DO PROCEDIMENTO PARA ANÁLISE DE BIG DATA .....	29
3.6. APLICAÇÃO DO PROCEDIMENTO NA BASE DE DADOS ESCOLHIDA (CRIPTOMOEDAS) .....	29
3.7. ANÁLISE DOS RESULTADOS .....	30
<b>4. PROCEDIMENTO PARA ANÁLISE DE BIG DATA</b> .....	31
4.1. IDENTIFICAÇÃO DA BASE DE DADOS .....	32
4.2. PREPARAÇÃO DOS DADOS.....	34
4.2.1. Identificar variáveis de estudo .....	35
4.2.2. Missing data.....	36

4.2.3. Remover os outliers.....	36
4.3. ANÁLISE DE CORRELAÇÃO .....	37
4.4. ANÁLISE DESCRITIVA.....	38
4.5. REGRESSÃO LINEAR.....	39
<b>5. CONCLUSÃO .....</b>	<b>45</b>
<b>REFERÊNCIAS.....</b>	<b>46</b>
<b>APÊNDICE A – BASE DE DADOS .....</b>	<b>49</b>
<b>ANEXO A – TABELA F .....</b>	<b>51</b>

## 1. INTRODUÇÃO

Nos últimos anos, com os crescentes avanços nas tecnologias digitais, foi possível conectar diversos dispositivos e essa integração de componentes inteligentes entre vários sistemas fornece a base para o que hoje é chamado de indústria 4.0, e isso tem sido um grande desafio para as indústrias de fabricação tradicional, se adaptar a essas novas tecnologias (NEGRI; FUMAGALLI; MACCHI, 2017).

O uso dessas tecnologias vem causando grandes impactos nos sistemas produtivos industriais, sendo utilizado cada vez mais sensores e equipamentos conectados em rede para se ter um controle maior sobre a produção, tendo assim uma união do mundo real e virtual, juntamente com a utilização da inteligência artificial (IA) criando o que se chama de Sistemas Físico-Cibernéticos (CPS – Cyber-Physical Systems) (GOMES et al., 2016).

A tendência é que nos próximos anos essas tecnologias estejam presentes em todas as áreas da economia, juntamente a isso mais dispositivos conectados e capazes de se comunicarem, como smartphones, veículos, eletrodomésticos entre outros, somado ao processamento desses dados que serão gerados por eles, causem alterações nos modelos de negócios e também a forma com que as empresas se relacionam com os clientes (GOMES et al., 2016).

Diante da quantidade de dados gerados é de extrema importância que eles sejam analisados, com o intuito de se obter informações úteis, necessitando assim sistemas com uma boa capacidade de coletar, processar e gerenciar esses dados (SANTOS et al., 2017).

Dado o contexto da indústria 4.0 e a quantidade de dados que são gerados, tanto no ambiente interno das empresas quanto dados provenientes de fora, é de suma importância que as pessoas estejam preparadas para lidar com essas informações, por isso essa área vem ganhando espaço nos últimos anos e saber gerenciar essa grande quantidade de dados (Big Data) é importante para tomada de decisões das companhias. Uma grande fonte para se obter dados é o Blockchain, que é uma espécie de um grande registro de dados onde uma vez registrados é praticamente impossível de serem alterados, sendo assim uma importante ferramenta para extração de dados.

O blockchain é a ferramenta que está diretamente ligada às criptomoedas, pois elas utilizam desta tecnologia para terem os seus registros todos arquivados. Elas funcionam de maneira descentralizada, operando via diversas casas de câmbio em todos os locais do mundo, o qual gera uma diferença considerável no preço de uma mesma moeda em tempo real, essa diferença pode ser notada, mensurada e estudada analisando estatisticamente os dados, mas devida a tamanha quantidade de diferentes criptomoedas e da base de dados, é necessária uma análise estatística mais profunda de Big Data (MURRAY, 2013).

## 1.1. OBJETIVOS

Os objetivos foram divididos em objetivo geral e objetivos específicos, que são apresentados a seguir.

### 1.1.1. Objetivo geral

Propor um procedimento para análise de Big Data utilizando métodos estatísticos. O procedimento foi aplicado em uma base de dados em criptomoedas que é considerada um Big Data do Blockchain.

### 1.1.2. Objetivos específicos

- Identificar as bases de dados mais adequadas (acessíveis, disponíveis e apropriadas) para o estudo em Blockchain.
- Identificar os métodos estatísticos apropriados para análise do Big Data estudado.
- Aplicar as técnicas estatísticas escolhidas na base de dado de criptomoedas (Big Data em Blockchain).
- Analisar os resultados estatísticos das criptomoedas estudadas e seus comportamentos.

## 1.2. JUSTIFICATIVA

Diante desse cenário atual, com muitos avanços tecnológicos desenvolvidos e aplicados na área industrial, é de suma importância que as mesmas sejam bem compreendidas por profissionais ligados a esse setor, portanto o presente trabalho é uma oportunidade de estudar e aprender um pouco mais sobre essas tecnologias, como Big Data, internet das coisas (IoT), computação em nuvem, entre outras que são fundamentais para a indústria 4.0.

Dentro desse cenário, o surgimento da tecnologia do blockchain com o bitcoin, a primeira criptomoeda, revolucionou a forma como se tem a rastreabilidade dos dados, garantindo sua autenticidade e a imutabilidade das informações, com aplicações não somente nas criptomoedas, mas na indústria em geral.

Para realizar um estudo mais elaborado e aprofundado o foco deste trabalho será criptomoedas, tendo em vista que os seus pilares também são algumas das tecnologias como: Big Data, sistemas integrados, computação em nuvem, e juntamente com isso a facilidade para a obtenção da base de dados das moedas, que estão disponíveis na internet.

Devido a particularidade das criptomoedas serem negociadas em casas de câmbio independentes, seu preço variar de acordo com o volume negociado em um mesmo horário, a possibilidade de o movimento de uma moeda influenciar o movimento de as demais e a velocidade quase que instantânea que se transfere esses valores, foi notado uma oportunidade de negócio nessa área, analisando estatisticamente o comportamento das criptomoedas de modo a tentar agrupa-las por características semelhantes, identificar correlações entre elas e assim prever o movimento do mercado, com base no comportamento dessas moedas em diferentes casas de câmbio e suas influências.

## 2. FUNDAMENTAÇÃO TEÓRICA

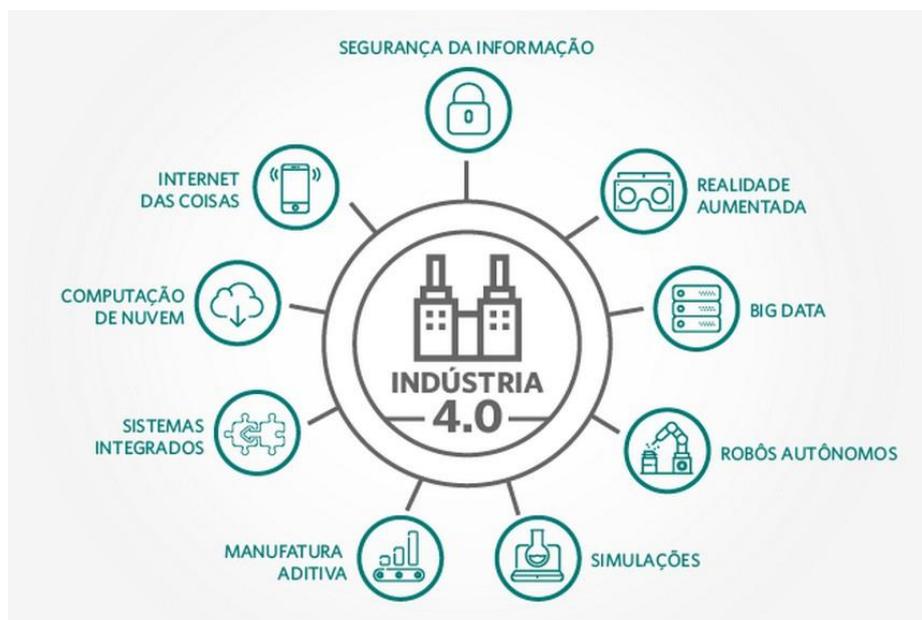
Neste capítulo será feito uma abordagem teórica sobre os temas relacionados a este trabalho.

### 2.1. INDÚSTRIA 4.0

O termo indústria 4.0 foi utilizado pela primeira vez na Alemanha em 2011, numa feira em Hanover, para se referir a esse novo conceito de “fábricas inteligentes” que estava surgindo (SCHWAB, 2016). Hoje em dia no ano de 2019, esse termo já é bastante utilizado tendo em vista que estamos passando por uma fase de transformações nas indústrias, muitos afirmam que estamos vivendo a quarta revolução industrial, no entanto ela não está relacionada apenas com sistemas industriais, seu escopo abrange mais temas que vão desde o sequenciamento genético até a nanotecnologia, das energias renováveis até a computação quântica. A fusão dessas tecnologias e a interação entre os domínios físicos, digitais e biológicos é o que diferencia a quarta revolução das anteriores (SCHWAB, 2016).

Para que essas “fábricas inteligentes” funcionem é fundamental que haja uma integração entre os fatores produtivos internos (planejamento da produção e outros setores da empresa como área de compras por exemplo) e também de fatores externos a empresa (fornecedores, demanda do mercado, entre outros), fazendo um trabalho em conjunto dos sistemas físicos e virtuais. Para isso é necessário a utilização de diversas tecnologias como: Big Data, IoT, inteligência artificial (IA), realidade aumentada, entre outras, conforme ilustrado na Figura 1, e assim é possível que as máquinas tomem suas próprias decisões sem que seja necessária uma intervenção humana (GERBERT et al., 2015).

**Figura 1 – Indústria 4.0 e suas tecnologias**



**Fonte: LWTSISTEMAS, 2019**

A indústria 4.0 irá causar vários impactos tanto para os clientes, que terão um controle maior sobre os produtos pois poderão customiza-los e personaliza-los, quanto para quem trabalha diretamente nas fábricas, uma vez que os profissionais terão que se adaptar aos novos sistemas automatizados e a tendência é que os trabalhos manuais e repetitivos sejam substituídos todos por mão de obra automatizada, porém apesar de esses cargos deixarem de existir outros novas demandas de cargos deverão surgir, mas provavelmente exigindo um conhecimento técnico maior também (KANGERMANN; WAHLSTER; HELBIG, 2013). As grandes vantagens disso é que as fábricas serão muito mais eficientes, com um poder de tomada de decisão mais rápido conforme as necessidades do mercado, tendo um controle maior sobre a produção e os estoques e com isso diminuindo os custos dos processos (GERBERT, 2015).

## 2.2. BIG DATA

A quantidade de dados produzidas no mundo atualmente é muito grande, basta pensarmos nas nossas atividades diárias, é muito comum utilizarmos as redes sociais e se pensarmos na quantidade de informações, fotos e vídeos que são postadas já podemos imaginar uma quantidade grande de dados, porém essa quantidade é produzida apenas por nós humanos, se pensarmos também no que é produzido por

máquinas e outros dispositivos que estão conectados esse número fica ainda maior. Segundo a IDC (International Data Corporation), empresa de pesquisa de mercado e consultoria, estima-se que até 2020 a quantidade de dados produzidos no mundo inteiro seja de 44 zettabytes (44 trilhões de gigabytes).

O termo Big Data surgiu no início da década de 1990, devido ao aumento muito grande no volume de dados produzidos juntamente com a tecnologia que se tinha na época, que não era tão avançada quanto hoje, com isso começaram os problemas com a gestão e manipulação desses dados, a solução encontrada foi ao invés de melhorar a capacidade de desempenho de uma máquina para a análise de todos esses dados, tentar dividir essa função em mais máquinas com capacidade menores e que pudessem trabalhar ao mesmo tempo e em conjunto, ou seja, fazer o processamento desses dados em paralelo utilizando máquinas mais baratas. A definição de Big Data segundo a IDC:

“Big Data é uma nova geração de tecnologias e arquiteturas, desenhadas de maneira econômica para extrair valor de grandes volumes de dados, provenientes de uma variedade de fontes, permitindo alta velocidade na captura, exploração e análise dos dados” (IDC, 2011, p.6).

Quando falamos em Big Data é importante destacar os 5 V's, demonstrados na Figura 2, pois são 5 propriedades que nos ajudam a obter informações úteis desses dados. Conforme Rosangela Marquesone (2016) os 5 V's são descritos como:

- **Volume:** Como já foi citado, é o que impulsionou a criação do Big Data.
- **Variedade:** Os dados recebidos nem sempre estarão da mesma forma, eles podem ser obtidos de diversas fontes como um documento, um arquivo log, dispositivos, sensores de medição, entre outros, portanto antes de serem armazenados é importante saber se são dados estruturados, que possuem um formato definido como tabelas por exemplo, semiestruturados, que possuem formato pré-definido porém não com tanto rigor como os estruturados, e por último os não estruturados, que podem ser imagens e vídeos por exemplo.
- **Velocidade:** Além do volume e da variedade a velocidade com que os dados são gerados e processados também é importante.
- **Veracidade:** É necessário saber se os dados são confiáveis ou não, devido ao grande volume e também variedade é possível que existam dados inconsistentes.
- **Valor:** É importante saber quais dados são mais significativos para que sejam analisados com prioridade.

**Figura 2 – Os 5V's do Big Data**

**Fonte: CULTURA ANALÍTICA, 2019**

Pensando em negócios o Big Data tem grandes aplicações e em vários seguimentos de mercado, segundo Rosangela Marquesone (2016) as redes de entretenimento Netflix e Spotify, o primeiro para filmes e o segundo para músicas, são exemplos onde é usado para analisar o que os usuários assistiram e escutaram ultimamente e então dar recomendações a eles relacionadas. Nas indústrias com as “fábricas inteligentes” que tem um papel fundamental de analisar o mercado e a demanda do que está sendo produzido e assim saber se aumenta ou diminui a produção. No caso do setor financeiro é utilizado para verificar se os clientes estão satisfeitos ou não com os serviços e assim diminuir a taxa com que eles mudam suas contas para concorrentes, outra aplicação do Big Data é relacionada às criptomoedas (HEKIMA, 2015).

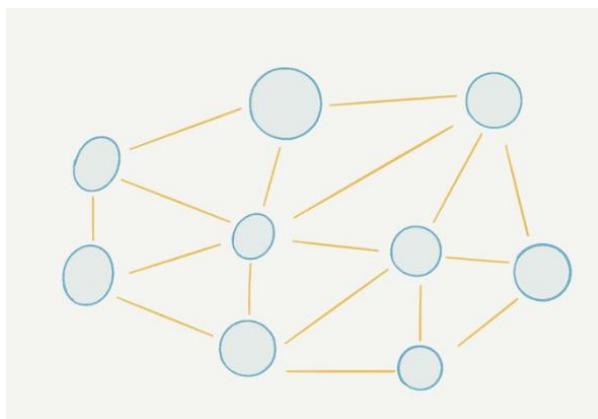
### 2.3. BLOCKCHAIN

Para entender o conceito de criptomoeda é necessário entender o que é o blockchain, a tecnologia base por trás das criptomoedas. Segundo O’Byrne (2016), O Blockchain (Corrente de blocos) é um livro de registro público descentralizado e distribuído em uma rede Ponto-a-Ponto (Peer to Peer).

Diferente do que se imagina, o blockchain não surgiu em 2008 junto com o bitcoin. A idéia já vem desde 1991 com o trabalho “How to time-stamp a digital document” de Scott Stornetta e Stuart Haber, onde estavam discutindo sobre como ter certeza sobre o passado e proteger isso para as gerações futuras sem depender de uma autoridade central. Nesse trabalho surge o conceito de blockchain, Hash e de encriptar essa informação de modo a garantir que não existam fraudes na rede (WHITAKER, 2018).

Uma rede Ponto-a-Ponto, Peer-to-Peer mostrado na Figura 3, é uma rede de computadores e servidores onde compartilham arquivos ou tarefas entre esses pares (Peer). Cada par possui privilégios idênticos dentro da rede e cada usuário, computador ou servidor dessa rede é chamado de nó. Quando uma nova informação entra nessa rede, ela se propaga entre todos os nós, e sendo ela encriptada, é impossível determinar qual nó inseriu a informação na rede. Como a rede é descentralizada, caso um ou vários nós deixem a rede, todos os outros nós dessa mesma rede possuem exatamente a mesma informação e quando um novo nó se junta, toda essa informação é copiada para esse novo nó, o tornando parte da rede (O'BYRNE, 2016).

**Figura 3 - Esquema de rede ponto-a-ponto (peer-to-peer)**



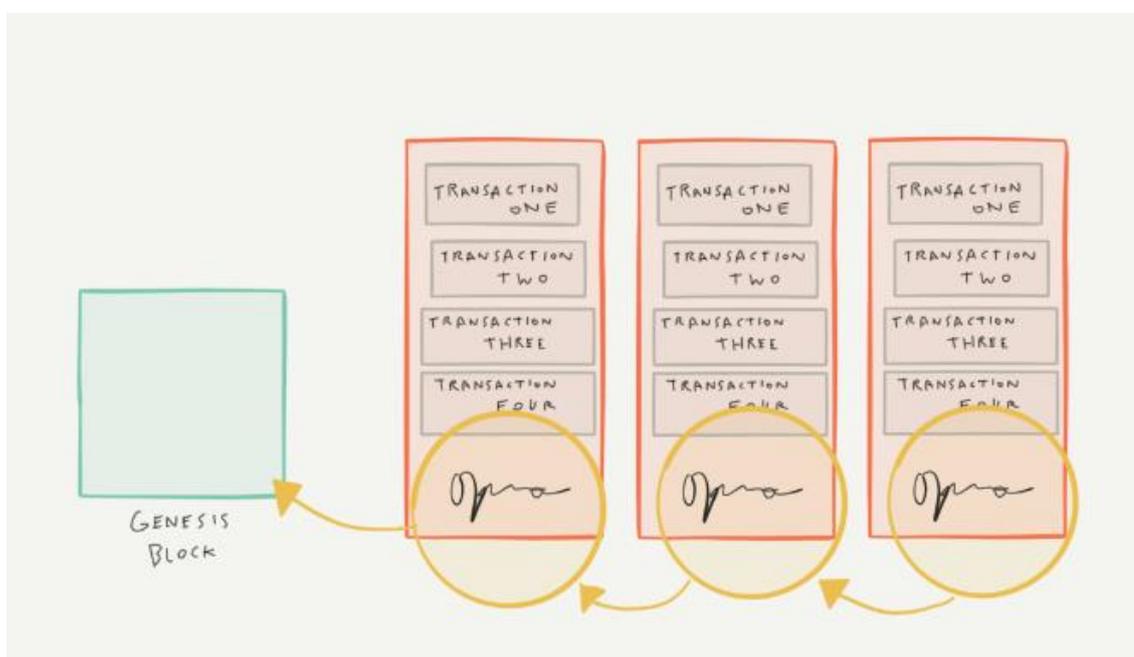
**Fonte: O'BYRNE, 2016**

A informação propagada no Blockchain, são as transações feitas na rede e o primeiro elemento do banco de dados é chamado de Block Genesis (Genesis Block), que contém as informações sobre as regras e direções de todo a base de dados. A partir desse bloco, a base de dados é formada criando-se novos blocos que irão conter as informações ou os registros dessas transações e serão gravadas e alocadas no bloco de acordo com a hora que foram criadas. Cada bloco é conectado ao anterior

formando uma série de blocos da qual vem o nome da tecnologia, Blockchain (corrente de blocos) (O'BYRNE, 2016).

A combinação da informação e o momento em que foi criada no bloco faz-se o livro registro. Quando uma série de informações são agrupadas no bloco, é criada uma assinatura criptografada chamada Hash no fundo desse bloco e essa assinatura é a responsável por linkar o bloco atual ao bloco anterior. Cada assinatura possui um número, que caracteriza o número atual do bloco e também o número do próximo bloco, assim como a data e hora em que foi criada e quantas transações existe no bloco. Seguindo as assinaturas, é possível ordenar todos os blocos levando do final até o bloco Genesis, esquema demonstrado na Figura 4.

**Figura 4 - Estrutura do Blockchain**



**Fonte: O'BYRNE, 2016**

Conforme novos blocos vão sendo criados, eles recebem sua assinatura e são adicionados a essa corrente, assim formando uma corrente de blocos (Blockchain).

## 2.4. CRIPTOMOEDAS

Segundo Greenberg (2011), Criptomoeda é um tipo de moeda digital que se utiliza da tecnologia do blockchain e criptografia para assegurar a validade de elas transações assim como a criação de novas unidades de moeda. O surgimento das

criptomoedas se dá em 2008 com o White Paper, um documento oficial que aprofunda um determinado problema, descrevendo suas causas e sua solução (KOLOWICH, 2018), foi feito pelo programador ou grupo de programadores não identificado chamado Satoshi Nakamoto, onde ele cria um sistema de transações online descentralizado usando a tecnologia Peer-to-peer e o Blockchain e que utiliza como moeda de troca, uma moeda digital gerada na própria rede chamada Bitcoin (ULRICH, 2014).

O Bitcoin, segundo Antonopoulos, (2016) “É um conjunto de conceitos e tecnologias que forma a base de um ecossistema de dinheiro digital” onde, diferentemente das moedas tradicionais, é uma moeda totalmente virtual. Diversas moedas digitais foram criadas ao longo dos anos, principalmente no final dos anos 80 com o avanço da criptografia. Esses projetos emitiam dinheiro virtual e eram lastreados com ouro ou moedas tradicionais, porém acabavam enfrentando 2 problemas comuns nesse tipo de projeto: O de segurança do sistema e de como evitar gastos duplicados. Para evitar os gastos duplicados, essas moedas digitais usavam uma central de compensação para finalizar todas as transações em intervalos pré-determinados, como é feito no sistema bancário tradicional. Devido a essa centralização, a segurança do sistema era comprometida e ficando não só vulnerável a ataques de hackers, mas também sob o controle de uma entidade centralizadora e controladora de todo o sistema (ANTONOPOULOS, 2016).

A grande inovação do Bitcoin foi resolver com a computação distribuída o problema da centralização da moeda e junto com o blockchain, resolver o problema dos gastos duplicados. Os Bitcoins são gerados através de um processo chamado de mineração, que consistem em resolver um problema matemático enquanto se processam as transações de Bitcoin na rede (ANTONOPOULOS, 2016). Esses problemas matemáticos, são blocos criptografados criados pela rede Bitcoin, que aproximadamente seis vezes por dia, um lote com no máximo 50BTC é criado e distribuídos aleatoriamente entre os participantes da rede que estão rodando o programa de mineração (NAKAMOTO; SATOSHI, 2009). Todos os mineradores disputam a resolução desse bloco, e quando um nó minerador acha a solução, ele diz a rede que achou a solução, a rede valida a solução, o prêmio em Bitcoin é reivindicado a carteira vencedora e o bloco resolvido é adicionado ao blockchain. O processo de mineração, descentraliza a função de compensação e emissão da

moeda, que é atribuída a um banco central, eliminando a necessidade de um órgão emissor e centralizador de moeda.

Esse processo de mineração, desenvolvido por Nakamoto, é baseado no conceito de prova de trabalho (proof-of-work), onde é usada para atingir o consenso da rede sem a necessidade de uma autoridade central confiável. Esse é um grande avanço na ciência da computação distribuída e tem amplas aplicações não somente na criação de sistemas de pagamento, mas ele pode alcançar o consenso em redes descentralizadas para provar honestidade de eleições, registro de bens, loterias, certificações e mais (ANTONOPOULOS, 2016).

## 2.5. ESTATÍSTICA DESCRITIVA

O objetivo da estatística descritiva é dar uma ideia geral sobre o conjunto de dados que está sendo analisado. Para auxiliar nessa visualização geral sobre os dados alguns parâmetros são calculados como a média, mediana, variância e desvio padrão. Os dois primeiros são classificados como medidas de tendência central, porque buscam sempre um valor que esteja mais centralizado no grupo de valores observado, já os dois últimos são classificados como medidas de dispersão, que ajudam as medidas de tendência central para explicar como é o conjunto estudado, indicando se os dados estão próximos ou não. Por exemplo imagine um conjunto de amostras que tenha um desvio padrão baixo, significa que o grupo inteiro tem valores próximos a média, ou se o desvio padrão for nulo significa que todas têm o mesmo valor da média. Por outro lado, se o valor do desvio padrão for muito alto isso quer dizer que a média não é um valor representativo para todo o grupo (REIS, 2008).

### 2.5.1. Média aritmética

A média aritmética de um conjunto nada mais é do que a soma de todos os valores da variável estudada dividida pelo número total de observações, ou seja, em um conjunto com  $n$  pontos a média será a soma de todos os valores dos  $n$  pontos dividido por  $n$  (REIS, 2008).

### 2.5.2. Mediana

A mediana é uma medida alternativa à média aritmética que tem por objetivo indicar o valor que está mais ao centro do conjunto estudado, tendo assim metade dos valores a cima e outra metade a baixo da mediana, ou seja, um grupo com  $n$  amostras dispostas com valores em ordem crescente a amostra que estiver mais ao centro é a mediana. Por exemplo um conjunto com 5 números [ 4; 2; 7; 1; 9 ] a mediana será o número 4, basta colocar os valores em ordem crescente para perceber que o valor central será o 4. No caso onde dois valores tiverem ao centro é necessário calcular a média entre os dois valores, por exemplo um grupo com 6 números [ 1; 2; 2; 4; 6; 9 ] a mediana será 3, pois é a média entre 2 e 4, valores centrais do intervalo (REIS, 2008).

### 2.5.3. Variância

Para o cálculo da variância é levado em consideração o desvio de um ponto até a média, ou seja, o quanto este ponto está distante da média do conjunto em análise. Existe o cálculo da variância populacional e o cálculo da variância amostral (REIS, 2008).

Antes de continuar a explicação é necessário entender a diferença entre população e amostra. População é quando o conjunto está sendo analisado na sua totalidade, ou seja, o universo inteiro dessa variável está sendo considerado, já o termo amostra é para quando apenas uma parcela, um subgrupo da população está sendo analisado.

A variância é a soma dos quadrados das diferenças entre cada ponto e a média aritmética dividido por:  $N$  no caso da variância populacional Equação (1), e  $n - 1$  para a variância amostral Equação (2).

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} \quad (1)$$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad (2)$$

Onde os símbolos  $\mu$  e  $\bar{X}$  representam a média aritmética populacional e amostral respectivamente.

#### 2.5.4. Desvio padrão

Como o resultado da variância é um termo em unidades quadráticas isso dificulta o seu entendimento e sua interpretação, mas para isso temos o desvio padrão que nada mais é do que a raiz quadrada da variância, demonstrado pela Equação (3) para populacional e pela Equação (4) para a amostral.

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}} \quad (3)$$

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (4)$$

#### 2.6. ANÁLISE DE CORRELAÇÃO

Estudar o relacionamento entre duas ou mais variáveis é constantemente necessário, e o estudo desse relacionamento é chamado de estudo de correlação e regressão. A correlação linear fornece um número que representa o grau de relacionamento linear entre duas ou mais variáveis, esse número é chamado de coeficiente de Pearson representado pela letra  $r$  (VIALI, 2014), demonstrado pela Equação (5):

$$r = \frac{\sum (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \cdot \sum (Y_i - \bar{Y})^2}} \quad (5)$$

Sendo X e Y duas variáveis aleatórias, medidas no mínimo em escala de intervalo e que tenham uma distribuição conjunta normal bivariada (VIALI, 2014). O valor de  $r$  varia entre -1 a +1 onde uma correlação positiva significa uma correlação direta onde se um valor aumenta, o outro aumenta na mesma direção. Uma correlação negativa, significa que quando o valor de uma variável aumenta, a outra diminui e quando o  $r$  é igual a 0, significa que as variáveis não possuem nenhuma correlação linear, sendo necessária a investigação de sua correlação por outros métodos.

Na Tabela 1 é possível interpretar os graus de correlação de  $r$ .

**Tabela 1 – Tabela de interpretação da correlação**

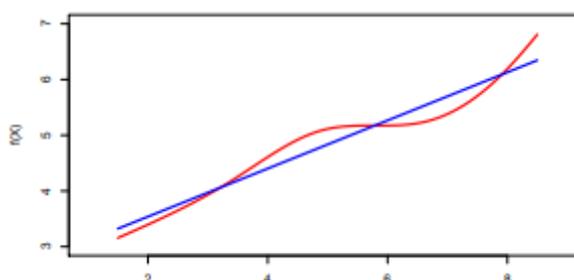
Valor de $r$ (- ou +)	Interpretação
0.00 a 0.19	Correlação bem fraca
0.20 a 0.39	Correlação fraca
0.40 a 0.69	Correlação moderada
0.70 a 0.89	Correlação forte
0.90 a 1.00	Correlação muito forte

**Fonte: SHIMAKURA, 2006**

A correlação linear pode ser simples ou múltipla, dependendo do número de variáveis a serem correlacionadas. Quando é correlacionado apenas 2 variáveis, é chamado de correlação simples, quando existem 3 ou mais variáveis a serem correlacionadas, o estudo é chamado de correlação múltipla (VIALI, 2014). Nosso estudo nesse trabalho irá apenas abordar correlações simples.

## 2.7. ANÁLISE DE REGRESSÃO LINEAR

Quando constatado que existe uma correlação linear entre duas variáveis, é possível tentar prever se o comportamento de uma das variáveis em função da outra é linear conforme a Figura 5 abaixo:

**Figura 5 – Relação linear entre X e Y**

**Fonte: STANFORD.EDU, 2019**

Para a regressão linear simples será suposto que existe apenas duas variáveis, sendo elas  $X$  e  $Y$ . A variável  $X$ , também chamada de variável explicativa, controlada ou independente, tem seus valores observados como  $X_1, X_2, X_3, \dots, X_n$  e a Variável  $Y$ , chamada de variável explicada ou dependente, tem seus valores como:  $Y_1, Y_2, Y_3, \dots, Y_n$ . Os valores de  $Y$  são aleatórios, não dependentes apenas  $X$  mas como

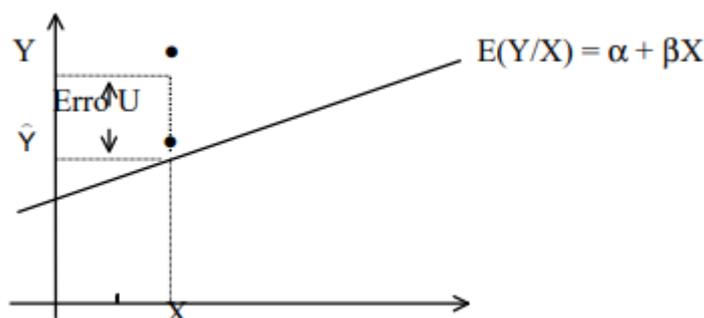
também de variáveis não consideradas pelo modelo. Essas variáveis não consideradas serão tratadas através de um termo aleatório chamado erro (VIALI, 2014).

A Equação 6 representa o modelo para esse relacionamento linear entre X e Y representado por:

$$Y = \alpha + \beta X + U \quad (6)$$

Onde U é o termo erro, o qual representa as influências exercidas na variável Y além da variável X. Sendo U de forma ideal, pequeno e independente de X, é possível modificar X sem modificar U e assim determinar o que ocorrerá na média em Y, assim podemos resumir:  $E(Y/X) = \alpha + \beta X$ . A Figura 6 abaixo ilustra graficamente a equação:

**Figura 6 – Gráfico da equação modelo**



**Fonte: VIALI, 2014**

Um modelo de regressão possui algumas hipóteses em relação ao termo erro e suas relações com as variáveis X e Y sendo duas delas:

$$E(U_i) = 0 \quad (7)$$

$$\text{Var}(U_i) = \sigma^2 \quad (8)$$

Onde na Equação 7 supõem-se que  $U_i$  são variáveis independentes e aleatórias com valor esperado de 0 e na Equação 8 que a variância de  $U_i$  é a mesma e igual  $\sigma^2$  para todos os valores de X (VIALI, 2014).

### 2.7.1. Método dos mínimos quadrados

O método mais utilizado para se ajustar uma linha entre as variáveis X e Y é o método dos mínimos quadrados, que procura encontrar o melhor ajuste minimizando o a soma dos quadrados das diferenças entre os dados observados e valor estimado representado por  $\phi = \sum_{i=1}^n E_i^2$ .

Sendo  $E_i = Y_i - a - bX_i$  e a e b estimadores de  $\alpha$  e  $\beta$ , é necessário diferenciar a expressão parcialmente em relação a A e B (VIALI, 2014). Simplificando em:

$$\sum Y_i = na + b \sum X_i \quad (9)$$

$$\sum X_i Y_i = a \sum X_i + b \sum X_i^2 \quad (10)$$

Essas expressões são chamadas de equações normais da regressão onde n é o número de pares de observação.

Ao dividir a Equação (9) por n e isolando-se a, temos (11):

$$a = \frac{\sum y_i}{n} - b \left( \frac{\sum X_i}{n} \right) = \bar{Y} - b\bar{X} \quad (11)$$

Substituindo esse resultado na Equação (10) temos (12):

$$b = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} \quad (12)$$

Assim podemos ter que a reta estimada da regressão é a Equação (13):

$$\hat{Y} = a + bX \quad (13)$$

### 2.7.2. Coeficiente de determinação $R^2$

O coeficiente de determinação é um indicador que mostra qual o percentual a variação do modelo de regressão representa da variação total.

O valor de  $R^2$  é um valor entre 0 e 1 onde se igual 1, significa que todos os pontos observados são exatamente sobre a reta de regressão, tendo um ajuste perfeito, significando também que 100% das variações de Y são explicadas pelas variações de X (VIALI, 2014)

O valor de  $R^2$  é calculado pela Equação (15):

$$R^2 = 1 - \frac{SQ_{res}}{SQ_{tot}} \quad (14)$$

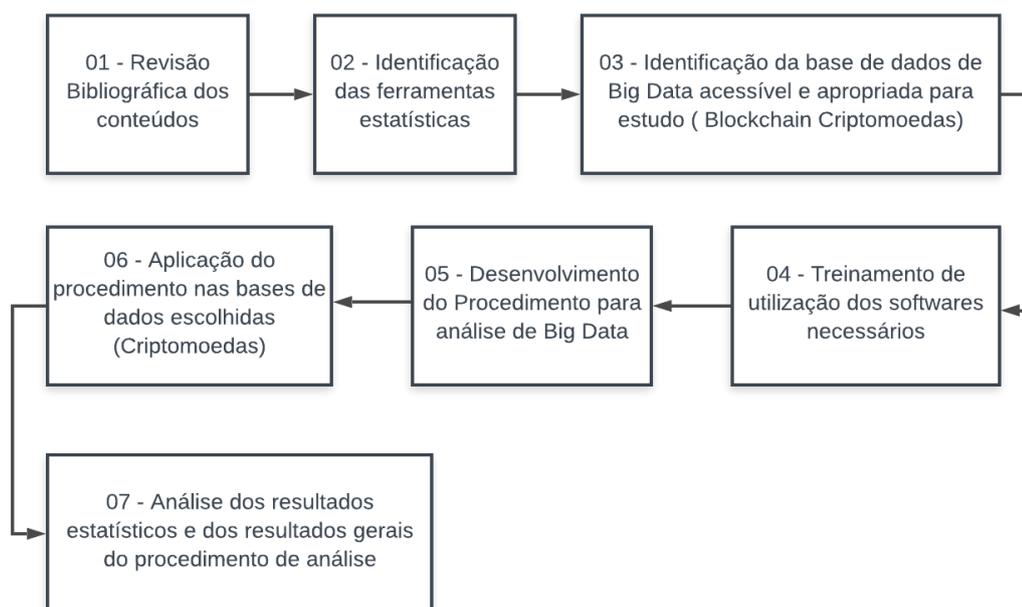
$$R^2 = 1 - \frac{\sum_{i=1}^N [y_i - (b \times x + a)]^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (15)$$

Onde  $SQ_{res}$  é o somatório dos quadrados dos resíduos e  $SQ_{tot}$  o somatório dos quadrados do total.

### 3. METODOLOGIA

A Figura 7 descreve o fluxograma da metodologia, onde apresenta sete passos para realizar o trabalho.

**Figura 7 – Fluxograma da metodologia**



**Fonte: Autoria própria**

Os sete passos descrito no fluxograma da metodologia (Figura 7) são detalhados em tópicos.

#### 3.1. REVISÃO BIBLIOGRÁFICA DOS CONTEÚDOS

Revisou toda a bibliografia dos conteúdos de estatísticas já citados nesse trabalho assim como obras de indústria 4.0, Blockchain e criptomoedas de modo a embasar o estudo proposto.

#### 3.2. IDENTIFICAÇÃO DAS FERRAMENTAS ESTATÍSTICAS

Foram identificadas as ferramentas estatísticas mais apropriada para a análise do problema de acordo com a aplicabilidade no problema proposto. As

ferramentas escolhidas foram o método descritivo, o método de análise de correlação e o método de análise de regressão linear.

### 3.3. IDENTIFICAÇÃO DA BASE DE DADOS DE BIG DATA ACESSÍVEL E APROPRIADA PARA O ESTUDO (BLOCKCHAIN – CRIPTOMOEDAS)

O estudo apresenta um procedimento de análise em Big Data, a base de Big Data utilizado foram as bases de criptomoedas, pois destaca a relação de interação entre as moedas, apresentadas nos Blockchain. As bases de criptomoedas foram escolhidas por serem bases de dados abertas e de fácil acesso. As bases foram extraídas de sites de base de dados específicos de criptomoedas.

### 3.4. TREINAMENTO DE UTILIZAÇÃO DOS SOFTWARES

Após a identificação das ferramentas estatísticas mais adequadas para a análise das bases de dados, identificou o software mais apropriado para realizar as análises, levando em conta o suporte do software ao formato da base de dados escolhida. O software escolhido foi Minitab19, uma ferramenta de análise estatística que atua no mercado a mais de 40 anos no mundo todo. A interface intuitiva e de fácil aprendizado foram essenciais para escolha da ferramenta.

### 3.5. DESENVOLVIMENTO DO PROCEDIMENTO PARA ANÁLISE DE BIG DATA

Após a identificação dos métodos estatísticos mais adequados para a análise da base de dados (big data) escolhida. Desenvolveu o procedimento para análise de big data. O procedimento de análise foi dividido em cinco etapas, apresentado os métodos estatísticos escolhidos em cada etapa.

### 3.6. APLICAÇÃO DO PROCEDIMENTO NA BASE DE DADOS ESCOLHIDA (CRIPTOMOEDAS)

Após o desenvolvimento do procedimento de análise em big data, foi realizado a aplicação do procedimento na base de dados escolhida, a base de criptomoedas, pois é uma base aberta e acessível.

### 3.7. ANÁLISE DOS RESULTADOS

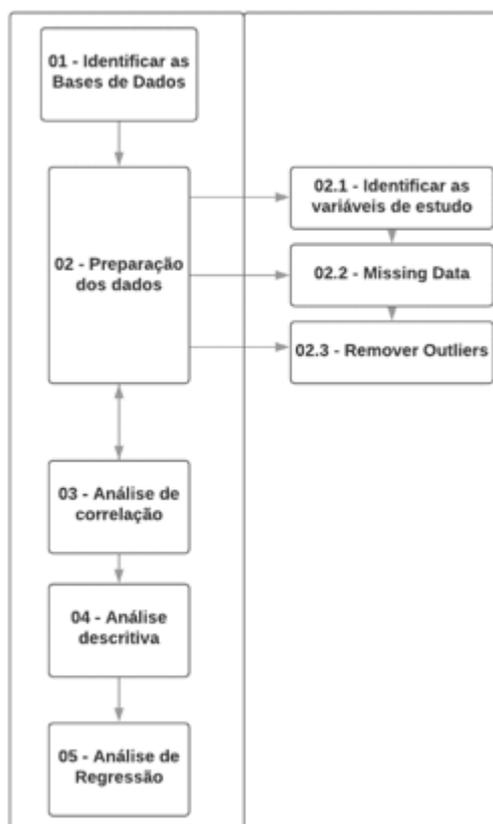
Aplicou o procedimento de análise desenvolvido na base de dados escolhidas, onde os principais resultados foram por meio da análise da estatística descritiva, da análise da correlação e da análise da regressão linear.

O procedimento para análise proposto foi apresentado no tópico 4 deste trabalho.

## 4. PROCEDIMENTO PARA ANÁLISE DE BIG DATA

A Figura 8 apresenta um fluxograma descrevendo o procedimento de análise proposta pelo trabalho.

**Figura 8 – Procedimento de análise de Big Data para Blockchain**



**Fonte: Autoria própria**

A Figura 8 descreve que o procedimento de análise foi dividido em cinco etapas principais. A primeira etapa identificou a base de dados estudadas, a base de dados foi a criptomoedas, pois é uma base de dados que tem interação e faz parte do conjunto do Blockchain.

A segunda etapa foi a preparação de dados onde foi realizado a identificação das variáveis de estudo, identificou que a variável estudada foi a variação das criptomoedas. Depois foi realizado o estudo de Missing Data (Dados faltantes), onde foram substituídos os dados faltantes pelas médias das variáveis. Na preparação dos dados também foi realizado a remoção dos outliers, que eliminou os dados que foram considerados mais dispersos das amostras estudadas.

Na terceira etapa, foi realizado a análise de correlação que é um dos principais resultados do trabalho. A análise de correlação foi executada juntamente com a preparação, pois a correlação influenciou no estudo de Missing Data.

A análise descritiva foi a quarta etapa, onde foram identificados as médias, medianas e desvios padrões das amostras de cada variável (criptomoedas) estudadas.

E por fim foi realizado a análise de regressão linear, sendo esse um dos principais resultados do trabalho. A regressão linear identificou a relação entre as variáveis analisadas (criptomoedas) e suas influências.

#### 4.1. IDENTIFICAÇÃO DA BASE DE DADOS

Para a escolha da base de dados do processo algumas condições foram levadas em consideração. A primeira delas foi a utilização da tecnologia do blockchain, já mencionada anteriormente, pois é algo que vem ganhando espaço dentro das grandes companhias e é importante para o sucesso das empresas na indústria 4.0. O segundo fator avaliado foi quanto a estratégia para obter uma base de dados com uma quantidade considerável, porque analisou-se big data, sendo assim essencial que a base de dados seja grande para que fosse realizado um trabalho com análises compatíveis com o foco do estudo. Tendo isso em mente a alternativa encontrada foi utilizar as criptomoedas, pois um dos seus pilares e motivo do seu sucesso é justamente a utilização do blockchain, quanto a obtenção de dados é possível encontra-los tanto em sites especializados em criptomoedas ou em casas de câmbio e gratuitamente, o que foi um fator que pesou na escolha também, porque dificilmente teríamos acesso a alguma base de dados grande sem custos.

Com isso definido no próximo passo foi feita a coleta dos dados, utilizando o site especializado da Coin Market Cap (2019) foram escolhidas as 11 criptomoedas mais negociadas no mundo, no momento da pesquisa 21 de setembro de 2019, que estão demonstradas na Tabela 2. Pensando em fazer uma boa análise foram coletados os dados referentes ao período de 2 anos de cada moeda.

**Tabela 2 – Ranking das criptomoedas. Valores referentes a 21/09/2019**

Ranking	Moeda	Sigla	Total de Moedas Circulantes	Preço	Valor de mercado
1	Bitcoin	BTC	18.015.075	<b>R\$36.566,00</b>	R\$658.739.229.392
2	Ethereum	ETH	108.310.999	<b>R\$714,12</b>	R\$77.346.962.939
3	XRP	XRP	43.248.091.671	<b>R\$1,18</b>	R\$51.055.934.377
4	Bitcoin Cash	BCH	18.079.438	<b>R\$993,74</b>	R\$17.966.280.817
5	Litecoin	LTC	63.549.742	<b>R\$226,01</b>	R\$14.362.909.256
6	EOS	EOS	937.579.414	<b>R\$12,47</b>	R\$11.692.411.374
7	Binance Coin	BNB	155.536.713	<b>R\$74,59</b>	R\$11.601.560.538
8	Stellar	XLM	20.039.779.054	<b>R\$0,248834</b>	R\$4.986.570.379
9	TRON	TRX	66.682.072.191	<b>R\$0,064931</b>	R\$4.329.765.332
10	Cardano	ADA	25.927.070.538	<b>R\$0,159826</b>	R\$4.143.811.706
11	Monero	XMR	17.278.172	<b>R\$223,50</b>	R\$3.861.598.568

**Fonte: COIN MARKET CAP, 2019**

O ranking na Tabela 2 acima é referente ao ranking de valor total de mercado das moedas, o ranking total pode ser acessado no site Coin Market Cap onde o ranking total contém mais de 2mil moedas circulantes. A razão dessas criptomoedas serem as primeiras são variadas, sendo desde sua aplicabilidade, uma vez que cada moeda tem um propósito e uma utilidade específica, até razões puramente especulativas devido a relação custo benefício de seu valor de mercado e o custo de mineração. Para um melhor entendimento da Tabela 2 acima:

**Total de Moedas Circulantes:** É o número total de criptomoedas existentes no mercado, onde esse número pode crescer diariamente dependendo do tipo de origem da criptomoeda, se ela é uma moeda minerável ou não. No caso do bitcoin, uma criptomoeda minerável, esse número pode ser acessado diariamente em BLOCKCHAIN (2019). Moedas não mineráveis, possuem um número fixo de moedas circulantes no mercado sendo possível ao longo do tempo serem lançadas mais criptomoedas ao mercado ou não.

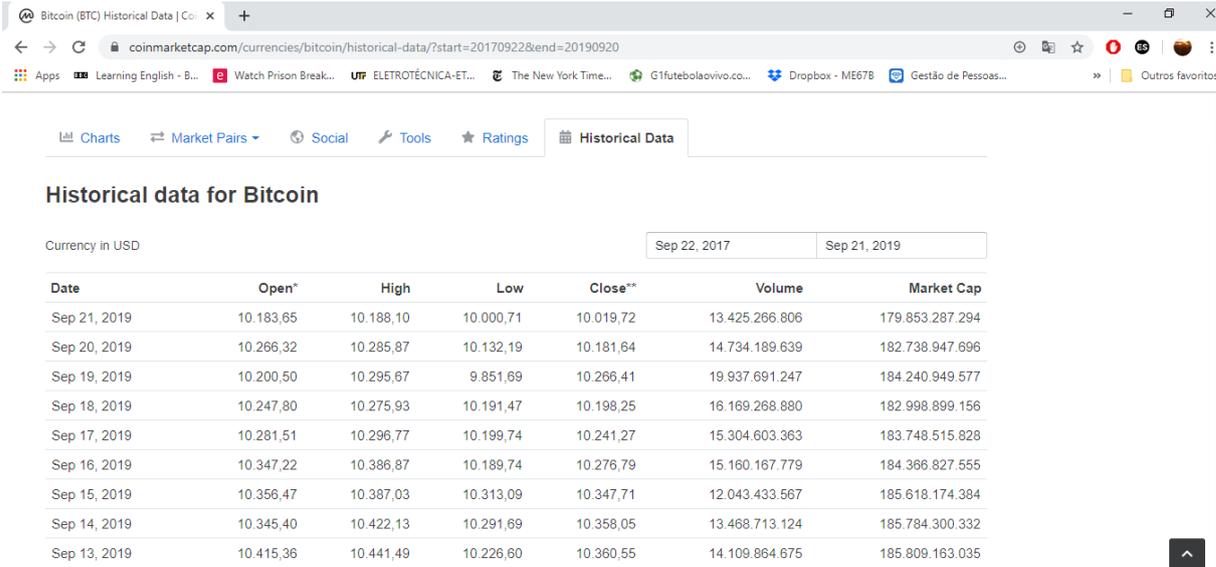
**Preço:** É o último valor de negociação de uma unidade de criptomoeda no mercado em BRL R\$.

**Valor de Mercado:** É o preço de uma unidade de criptomoeda multiplicado pelo total de moedas circulantes no mercado. Ela representa o valor total de mercado da criptomoeda e é um parâmetro usado de referência para relevância das criptomoedas.

## 4.2. PREPARAÇÃO DOS DADOS

O site da Coin Market Cap fornece os dados contendo seis variáveis para cada dia, conforme a Figura 9. As seis variáveis são: Open, High, Low, Close, Volume e Market Cap.

**Figura 9 – Fonte dos dados coin market cap**



Historical data for Bitcoin

Currency in USD

Sep 22, 2017    Sep 21, 2019

Date	Open*	High	Low	Close**	Volume	Market Cap
Sep 21, 2019	10.183,65	10.188,10	10.000,71	10.019,72	13.425.266.806	179.853.287.294
Sep 20, 2019	10.266,32	10.285,87	10.132,19	10.181,64	14.734.189.639	182.738.947.696
Sep 19, 2019	10.200,50	10.295,67	9.851,69	10.266,41	19.937.691.247	184.240.949.577
Sep 18, 2019	10.247,80	10.275,93	10.191,47	10.198,25	16.169.268.880	182.998.899.156
Sep 17, 2019	10.281,51	10.296,77	10.199,74	10.241,27	15.304.603.363	183.748.515.828
Sep 16, 2019	10.347,22	10.386,87	10.189,74	10.276,79	15.160.167.779	184.366.827.555
Sep 15, 2019	10.356,47	10.387,03	10.313,09	10.347,71	12.043.433.567	185.618.174.384
Sep 14, 2019	10.345,40	10.422,13	10.291,69	10.358,05	13.468.713.124	185.784.300.332
Sep 13, 2019	10.415,36	10.441,49	10.226,60	10.360,55	14.109.864.675	185.809.163.035

**Fonte: COIN MARKET CAP, 2019**

Para uma melhor compreensão da Figura 9:

01-Open: Primeiro preço registrado em uma negociação no dia.

02- High: Maior preço registrado em uma negociação no dia.

03- Low: Menor preço registrado em uma negociação no dia.

04- Close: Último preço registrado em uma negociação no dia.

05- Volume: Valor total negociado no dia.

06- Market Cap: Valor total da soma de todas as moedas disponíveis multiplicado pelo preço de fechamento do dia (Equação 16).

$$\text{Total de moedas} \times \text{Close} \quad (16)$$

Algumas observações são necessárias:

- Valor na base de dados é em Dólar americano USD.

- Market Cap é o valor de todas as moedas existentes e não do volume negociado no dia.

O preço de abertura (Open) e fechamento (Close) é uma referência ao preço final de início e abertura no dia. No caso das criptomoedas, como são negociadas 24/4, essa referência é usada pelo Coin Market Cap como: Open: 00:01:00 UTC Time e Close 23:59:99 UTC Time, ou seja, por isso existem diferenças nos preços de abertura de um dia com o de fechamento do dia anterior, porque existe um pequeno espaço de tempo onde isso não é contabilizado.

Os dados foram coletados e armazenados em uma planilha no Excel para que assim ficassem todas as 11 Criptomoedas juntas, com as linhas do programa indicando a data de cada uma delas. Ao final dessa etapa tendo 730 linhas (365 dias por ano) para cada moeda, chegamos em 8020 linhas (11 criptomoedas\*730 linhas = 8030 linhas), onde 10 faltantes foram da Cardano (explicado mais a frente), cada linha continha 6 dados, ou seja, 48120 dados foram analisados ao todo.

#### 4.2.1. Identificar variáveis de estudo

Com os dados todos dispostos no Excel foi necessário identificar qual seria a variável a ser estudada, a escolhida foi a variação do preço no dia, entretanto esse não era um dado fornecido diretamente pelo Coin Market Cap e para isso foi necessário fazer um tratamento de dados, adicionando uma coluna para esta informação, mostrado na Tabela 3, que foi preenchida através do seguinte cálculo:

$1 - \left(\frac{Close}{Open}\right)$  tendo assim a variação percentual do preço no dia.

**Tabela 3 – Nova coluna com a variação**

Moeda	Date	Open*	High	Low	Close**	Volume	Market Cap	Variação
Bitcoin	Sep 21, 2019	10.183,65	10.188,10	10.000,71	10.019,72	13.425.266.806	179.853.287.294	1,61%
Bitcoin	Sep 20, 2019	10.266,32	10.285,87	10.132,19	10.181,64	14.734.189.639	182.738.947.696	0,82%
Bitcoin	Sep 19, 2019	10.200,50	10.295,67	9.851,69	10.266,41	19.937.691.247	184.240.949.577	-0,65%
Bitcoin	Sep 18, 2019	10.247,80	10.275,93	10.191,47	10.198,25	16.169.268.880	182.998.899.156	0,48%

**Fonte: Autoria própria**

Devido a particularidades da ferramenta usada nesta análise, o Minitab, a base foi remodelada de forma que o software conseguisse trabalhar. Assim, os dados foram reorganizados em uma nova planilha contendo apenas as datas e a variável a ser estudada para cada uma das 11 criptomoedas, conforme a Tabela 4 abaixo:

**Tabela 4 – Nova planilha**

Date	Bitcoin	Ethereum	XRP	Bitcoin Ca	Litecoin	EOS	Binance Ci	Stellar	Cardano	Monero	TRON
Sep 21, 2019	1,61%	1,09%	1,70%	1,03%	2,21%	0,00%	1,40%	3,76%	0,79%	1,58%	-4,78%
Sep 20, 2019	0,82%	1,37%	2,64%	2,61%	2,35%	-0,75%	1,20%	9,23%	0,29%	2,73%	1,41%
Sep 19, 2019	-0,65%	-4,75%	3,60%	0,29%	1,10%	2,69%	2,83%	2,10%	1,49%	7,01%	1,98%
Sep 18, 2019	0,48%	-1,61%	-10,42%	-1,67%	-5,04%	-0,25%	-5,95%	-30,59%	-8,87%	-10,20%	-7,31%
Sep 17, 2019	0,39%	-5,83%	-9,26%	-3,89%	-2,25%	0,00%	-2,99%	-8,71%	-4,03%	1,41%	-6,98%

**Fonte: Autoria própria**

Uma parte do começo e uma parte do final dessa planilha se encontra no final, no Apêndice A.

#### 4.2.2. Missing data

Como a Criptmoeda chamada Cardano estava com dados faltantes, porque seus dados foram fornecidos a partir do dia 02 de outubro de 2017, portanto faltando 10 dias para completar os dois anos, foi necessário fazer um novo tratamento de dados e a escolha para este caso foi completa-los com a média das suas variações percentuais dos preços nos outros dias, conforme podemos observar na Tabela 5 e também no final no Apêndice.

**Tabela 5 – Completando os dados faltantes**

Date	Bitcoin	Ethereum	XRP	Bitcoin Ca	Litecoin	EOS	Binance Ci	Stellar	Cardano	Monero	TRON
Oct 01, 2017	-1,44%	-0,26%	-4,51%	4,21%	1,09%	1,94%	-19,53%	0,90%	<b>-0,49%</b>	2,48%	18,27%
Sep 30, 2017	-4,14%	-3,42%	-1,91%	0,92%	-4,77%	-0,41%	-0,79%	-5,29%	<b>-0,49%</b>	-1,40%	-14,02%
Sep 29, 2017	0,20%	2,56%	1,27%	2,43%	2,75%	12,20%	7,30%	4,43%	<b>-0,49%</b>	2,88%	31,64%
Sep 28, 2017	0,53%	2,39%	3,16%	1,95%	3,77%	-28,87%	-20,00%	3,16%	<b>-0,49%</b>	3,70%	-13,38%
Sep 27, 2017	-7,90%	-6,78%	-9,72%	-3,50%	-9,11%	-13,00%	-33,22%	-16,90%	<b>-0,49%</b>	-8,94%	-47,52%
Sep 26, 2017	0,92%	1,68%	-2,28%	0,79%	0,27%	1,73%	6,33%	-2,58%	<b>-0,49%</b>	0,72%	2,19%
Sep 25, 2017	-6,64%	-3,58%	-3,40%	-6,23%	-8,32%	-4,44%	-16,94%	-4,52%	<b>-0,49%</b>	-4,78%	21,06%
Sep 24, 2017	2,98%	1,30%	1,34%	1,81%	3,31%	5,85%	-0,96%	2,95%	<b>-0,49%</b>	2,46%	4,26%
Sep 23, 2017	-4,48%	-8,15%	-2,60%	-3,18%	-2,30%	-1,03%	-7,50%	-10,61%	<b>-0,49%</b>	-4,34%	-21,30%
Sep 22, 2017	-0,07%	-2,40%	-1,13%	0,68%	-3,09%	-4,04%	-2,44%	-0,72%	<b>-0,49%</b>	-2,32%	6,16%

**Fonte: Autoria própria**

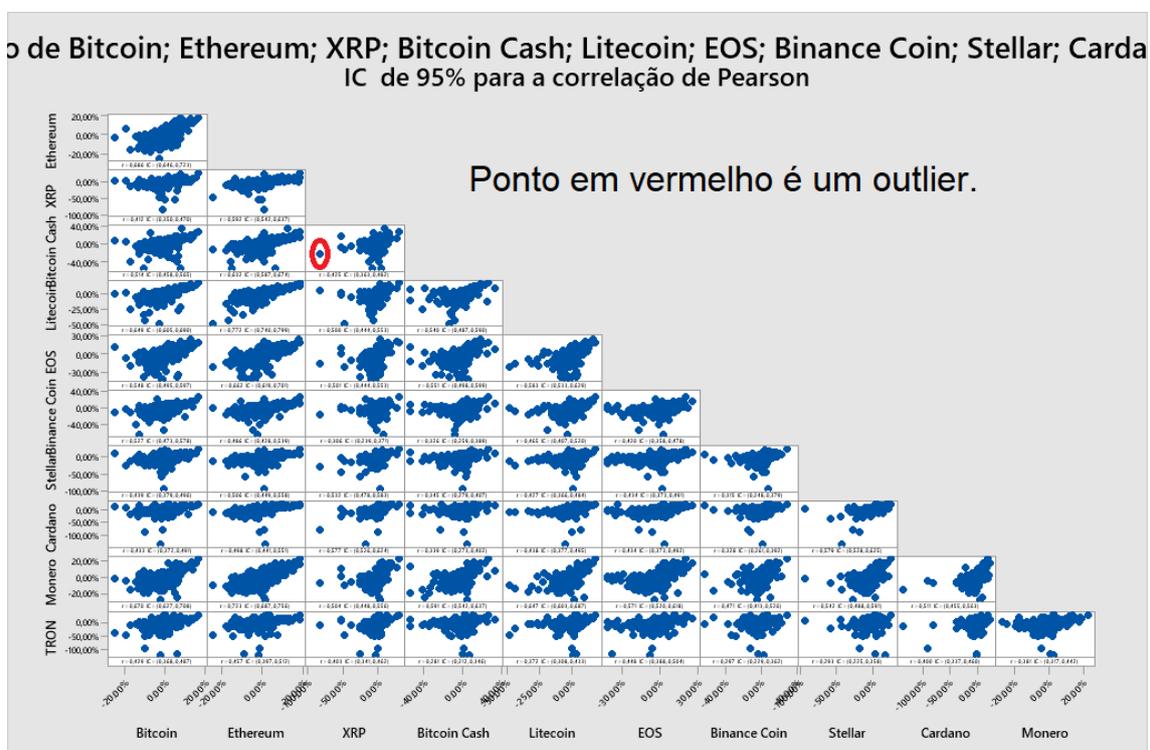
Agora com a base de dados pronta o próximo passo foi abrir o Minitab e importar essa planilha de dados do Excel para o software de estatística para então começar a trabalhar com eles.

#### 4.2.3. Remover os outliers

Esta parte do processo ocorreu em simultaneidade com o próximo passo, que é a análise de correlações, pois conforme estas eram feitas no Software e os gráficos

eram gerados, Figura 10, os pontos que estavam mais afastados, chamados de outliers, eram retirados de forma que a linha toda deste ponto era excluída.

**Figura 10 – Identificação de Outliers**



**Fonte: Autoria própria**

Conforme esses dados foram sendo retirados os valores das correlações foram aumentando até que se atingisse, na grande maioria, o critério estabelecido de 0,4 para as correlações, valor considerado moderado. No total foram excluídos 10 pontos, outliers.

### 4.3. ANÁLISE DE CORRELAÇÃO

Nesta etapa do processo foram realizadas no Minitab as correlações, utilizando o método de Pearson com um nível de confiança de 95%. Em uma primeira análise, ainda com todos os pontos, o resultado obtido foi o demonstrado na Tabela 6. Como podemos observar todas as correlações foram diretas, positivas (+), como já era esperado, porém com o intuito de aumentar esses valores alguns dados foram sendo retirados da análise, conforme descrito anteriormente, até que fosse atingido uma correlação moderada, pelo menos para a grande maioria.

**Tabela 6 – Correlação sem retirar os Outliers**

	Bitcoin	Ethereum	XRP	Bitcoin Cash	Litecoin	EOS	Binanc Coin	Stellar	Cardano	Monero
Ethereum	0,686									
XRP	0,412	0,592								
Bitcoin Cash	0,514	0,632	0,425							
Litecoin	0,649	0,772	0,5	0,54						
EOS	0,548	0,662	0,501	0,551	0,583					
Binance Coin	0,527	0,486	0,306	0,326	0,465	0,42				
Stellar	0,439	0,506	0,532	0,345	0,427	0,434	0,315			
Cardano	0,431	0,496	0,576	0,339	0,436	0,428	0,318	0,576		
Monero	0,67	0,723	0,504	0,591	0,647	0,571	0,471	0,542	0,509	
TRON	0,429	0,457	0,403	0,281	0,372	0,448	0,297	0,293	0,389	0,381

**Fonte: Autoria própria**

Como podemos observar na Tabela 7 os valores alcançados no final desta análise foram em sua grande maioria valores que estão dentro do intervalo moderado ( $0,4 < r < 0,69$ ) e alguns ainda ficaram no intervalo de correlação forte ( $0,7 < r < 0,89$ ).

**Tabela 7 – Correlação após a retirada dos Outliers**

	Bitcoin	Ethereum	XRP	Bitcoin Cash	Litecoin	EOS	Binanc Coin	Stellar	Cardano	Monero
Ethereum	0,732									
XRP	0,504	0,659								
Bitcoin Cash	0,554	0,64	0,432							
Litecoin	0,69	0,779	0,587	0,549						
EOS	0,609	0,67	0,512	0,559	0,603					
Binance Coin	0,561	0,525	0,357	0,335	0,486	0,455				
Stellar	0,521	0,605	0,618	0,394	0,501	0,509	0,428			
Cardano	0,601	0,699	0,621	0,45	0,611	0,576	0,465	0,676		
Monero	0,706	0,736	0,541	0,599	0,649	0,584	0,504	0,593	0,609	
TRON	0,514	0,585	0,517	0,414	0,54	0,505	0,43	0,468	0,49	0,517

**Fonte: Autoria própria**

Após a retirada dos pontos outliers foi realizado a análise descritiva para saber como estava a distribuição dos pontos.

#### 4.4. ANÁLISE DESCRITIVA

Com o auxílio do software Minitab foi realizada a etapa de análise descritiva o resultado é apresentado na Tabela 8.

**Tabela 8 – Resultados da análise descritiva**

Variável	N	N*	Média	DesvPad	Variância	Mediana
Bitcoin	720	0	-0,00141	0,04164	0,00173	-0,00152
Ethereum	720	0	-0,00076	0,05157	0,00266	0,0007
XRP	720	0	-0,00164	0,06607	0,00437	0,00216
Bitcoin Cash	720	0	-0,00224	0,07763	0,00603	0,00348
Litecoin	720	0	-0,00211	0,06092	0,00371	0,00311
EOS	720	0	-0,00446	0,07589	0,00576	0
Binance Coin	720	0	-0,00425	0,06449	0,00416	0,00086
Stellar	720	0	-0,00332	0,07269	0,00528	0,00231
Cardano	720	0	-0,00113	0,07048	0,00497	0,00135
Monero	720	0	-0,00072	0,05825	0,00339	0,00145
TRON	720	0	-0,00218	0,08421	0,00709	0,00341

**Fonte: Autoria própria**

Para um entendimento melhor da Tabela 8 podemos analisar os resultados mostrados da variável Bitcoin. A média aritmética dos 720 pontos da amostra foi -0,14%, ou como mostrado na imagem -0,00141. Através da Equação (2) foi calculado a variância amostral (124,664% dividido por 719) e o resultado foi 0,17%, valor mostrado na imagem 0,00173. E então utilizando a Equação (4) tem-se o valor do desvio padrão amostral que é 0,04164.

Analisando e comparando todos os resultados obtidos podemos perceber que o Bitcoin foi a variável com o menor valor de desvio padrão, ou seja, mostrando que era a amostra que tinha os valores mais próximos da média, enquanto que a criptomoeda TRON foi a que obteve o maior valor para o desvio padrão, mais que o dobro do Bitcoin, mostrando que os valores da amostra estavam mais dispersos.

#### 4.5. REGRESSÃO LINEAR

Após feito a análise de correlação uma análise de variância foi realizada junto com a regressão linear, sendo crucial sua interpretação para entendimento do resultado. Seu resultado é expresso em 5 valores estatísticos com valores referente a regressão, erro e o total conforme Tabela 9 abaixo e foi usado um grau de confiança de 0,95:

**Tabela 9 – Resultado variância Ethereum Bitcoin**

<b>Fonte</b>	<b>GL</b>	<b>SQ</b>	<b>QM</b>	<b>F</b>	<b>P</b>
Regressão	1	1,02582	1,02582	830,84	0,0000
Erro	718	0,8865	0,00123		
Total	719	1,91232			

**Fonte: Autoria própria**

GL: Grau de Liberdade – São os totais de informação nos dados. O GL total é o número de observações na amostra, o GL regressão é a quantidade de informação que cada componente usa, no nosso caso é apenas 1 e o GL Erro mostra quanta informação está disponível para se estimar a variabilidade de estimativas no parâmetro.

SQ: Soma dos Quadrados – São medidas da variação dos diferentes componentes do modelo e sua interpretação é:

SQ Regressão: A regressão da soma dos quadrados é a soma dos desvios quadrados dos valores de respostas ajustados do valor médio de resposta e mostra a quantidade da variação nos dados que são explicados pelo modelo proposto. O Minitab, usa esses valores para calcular o R2 Valor-P de um termo, que são os valores estatísticos de mais importante interpretação.

SQ Erro: é a soma dos resíduos quadrados e mostra a quantificação da variância nos dados que as preditoras não explicam.

SQ Total: Quantifica a variação total dos dados, é a soma do SQ Regressão e SQ erro.

QM: Quadrado médio – São valores quadrado médio ajustados e mostram quanta variação o modelo ou a amostra explica, independente mente da ordem em que os outros termos estejam no modelo. O QM diferentemente do SQ, levam em consideração os graus de liberdade do modelo. O QM, como o SQ, é usado para calcular o R2 e o Valor-P.

Valor F – é o teste usado para determinar se o modelo proposto está associado com a resposta. O Valor – F é usado para calcular o valor – P, e um valor amplo o suficiente indica que o modelo é significativo. Deve-se comparar o valor F calculado com o valor F crítico, um valor tabelado, e caso valor calculado seja maior, significa que o modelo tem significância, caso contrário pode-se rejeitar a hipótese. O valor F deve-se sempre ser usado em conjunto com o valor P para se tomar a decisão de rejeitar a hipótese. No Exemplo mostrado pela Tabela 9, temos um valor F calculado de 830,84 que deve ser comparado com o valor F crítico da tabela.

Para se achar o valor correto de F na tabela, deve-se procurar o valor crítico na tabela correta de acordo com o grau de confiabilidade usado no modelo, no nosso caso, 95% que significa procurar na tabela F de 5% de probabilidade. Nas colunas, deve-se usar o Grau de liberdade do modelo de regressão, no nosso caso 1 e nas linhas, o GL do erro, no nosso caso 718. Segundo a tabela que está em anexo para valores maiores que 240 e não infinitos, o valor crítico de F é de 3,881, sendo assim nosso valor F calculado da Tabela 9 muito maior. A tabela F de 5 % se encontra em sua totalidade no Anexo desse trabalho.

Valor P na regressão – Mede a evidência contra a hipótese nula, quanto menor o valor, mais forte é a evidência a hipótese nula. Sua interpretação: Para avaliar a hipótese nula, deve-se comparar o valor-p com o nível de significância do modelo. Hipótese nula para regressão, significa que nenhuma variação na resposta explica o modelo. Nível de significância  $\alpha$  no valor de 0,05 significa um risco de 5% que o modelo proposto explica a variação na resposta quando isso não acontece. Comparando o valor -p com  $\alpha$  temos que:

Se valor-p  $> \alpha$ : Significa que não há evidências suficientes para concluir que o modelo proposto justifica a variação na resposta.

Se valor-p  $\leq \alpha$ : Significa que o modelo explica a variação na resposta.

Na Figura 11 tem-se todos os resultados obtidos de análise de variância das amostras analisada.

Ao analisar o valor F da regressão de todos os casos, fica evidente que o valor F calculado é muito maior que o F crítico de 3,881 usado em todos os casos e em conjunto com o valor p menor que 0,05 proposto pelo modelo de 95% de confiabilidade, é possível afirmar que em todos os casos o modelo proposto é significativo.

**Figura 11 – Tabelas de resultados de análise de variância**

**Ethereum - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,02582	1,02582	830,84	0,0000
Erro	718	0,8865	0,00123		
Total	719	1,91232			

**XRP - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	0,79806	0,798065	244,81	0,0000
Erro	718	2,34062	0,00326		
Total	719	3,13868			

**Bitcoin Cash - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,32772	1,32772	317,24	0,0000
Erro	718	3,00503	0,00419		
Total	719	4,33275			

**EOS - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,53343	1,53343	422,3	0,0000
Erro	718	2,60717	0,00363		
Total	719	4,1406			

**Litecoin - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,27229	1,27229	654,18	0,0000
Erro	718	1,39642	0,00194		
Total	719	2,66871			

**Binance Coin**

Fonte	GL	SQ	QM	F	P
Regressão	1	0,94155	0,941552	330,01	0,0000
Erro	718	2,04851	0,002853		
Total	719	2,99007			

**Cardano - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,30221	1,30221	406,37	0,0000
Erro	708	2,2688	0,0032		
Total	709	3,57101			

**Stellar - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,03125	1,03125	267,5	0,0000
Erro	718	2,76796	0,00386		
Total	719	3,79921			

**Monero - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,21672	1,21672	714,23	0,0000
Erro	718	1,22315	0,0017		
Total	719	2,43987			

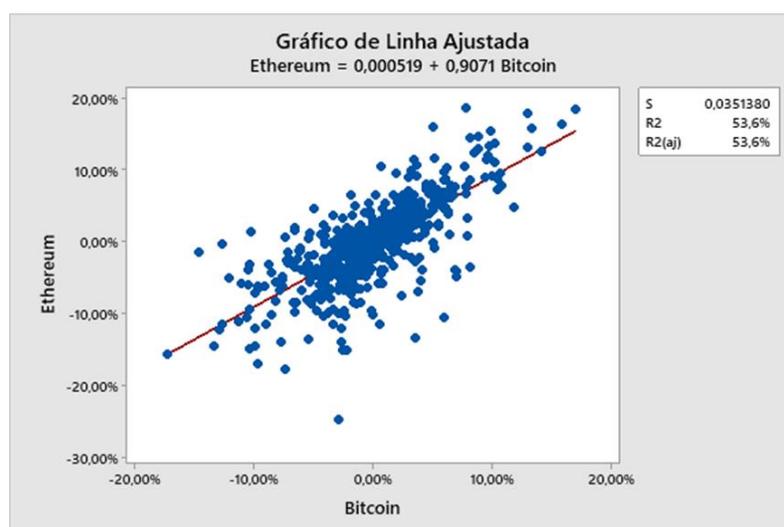
**TRON - Bitcoin**

Fonte	GL	SQ	QM	F	P
Regressão	1	1,34678	1,34678	257,72	0,0000
Erro	718	3,75209	0,00523		
Total	719	5,09888			

**Fonte: Autoria própria**

No sumário da análise de regressão linear, 3 informações importantes podem ser vistas: S, R2 e R2(aj) e o gráfico da linha ajustada é gerado como no exemplo da Figura 12 abaixo:

**Figura 12 – Gráfico de linha ajustado da moeda Ethereum**



**Fonte: Autoria própria**

A Figura 12 é um gráfico de linha ajustada referente a regressão do Bitcoin e Ethereum, onde mostra um gráfico de dispersão dos dados junto com a linha de regressão descrita pela equação de regressão, mostrada no título do gráfico.

A estatística S, primeiro valor mostrado no quadro de respostas, é o erro padrão da regressão, ou erro padrão da estatística e ele representa a distância média que os valores caem da linha de regressão. Em resumo, ele mostra o quanto o modelo está errando usando as unidades da variável resposta. Quanto menor o valor de S, mais próximas as observações estão da linha ajustada, que significa um melhor resultado. No gráfico exemplo, o valor S representa que em média, a distância dos dados da linha ajustada é de 3,5%.

Os valores de R e R(aj) são valores estatístico que representam o percentual de variação na resposta que é explicada pelo modelo, sendo o R(aj) ajustado para o número de preditores em relação ao número de observações. O valor de R2 e R2(aj) varia entre 0% a 100% e é usado para determinar o quão bem o modelo se ajusta aos dados, porém em caso de mais preditores, o R2(aj) é usado para ajudar a identificar qual o melhor modelo.

Após a análise feita no Minitab, tem-se como resposta as equações de linha ajustada assim como os valores de S, R2 e R2(aj) mostrados na Tabela 10:

**Tabela 10 – Resultados de regressão**

Moeda	Equação	S	R2	R2(aj)
Ethereum	Ethereum = 0,000519 + 0,9071 Bitcoin	0,0351	53,64%	53,58%
XRP	XRP = - 0,000507 + 0,8001 Bitcoin	0,0571	25,43%	25,32%
Bitcoin Cash	Bitcoin Cash = - 0,000786 + 1,032 Bitcoin	0,0647	30,64%	30,55%
Litecoin	Litecoin = - 0,000689 + 1,010 Bitcoin	0,0441	47,67%	47,60%
EOS	EOS = - 0,002899 + 1,109 Bitcoin	0,0603	37,03%	36,95%
Binance Coin	Binance Coin = - 0,003019 + 0,8691 Bitcoin	0,0534	31,49%	31,39%
Stellar	Stellar = - 0,002038 + 0,9095 Bitcoin	0,0621	27,14%	27,04%
Cardano	Cardano = 0,000109 + 1,028 Bitcoin	0,0566	36,47%	36,38%
Monero	Monero = 0,000672 + 0,9879 Bitcoin	0,0413	49,87%	49,80%
TRON	TRON = - 0,000708 + 1,039 Bitcoin	0,0723	26,41%	26,31%

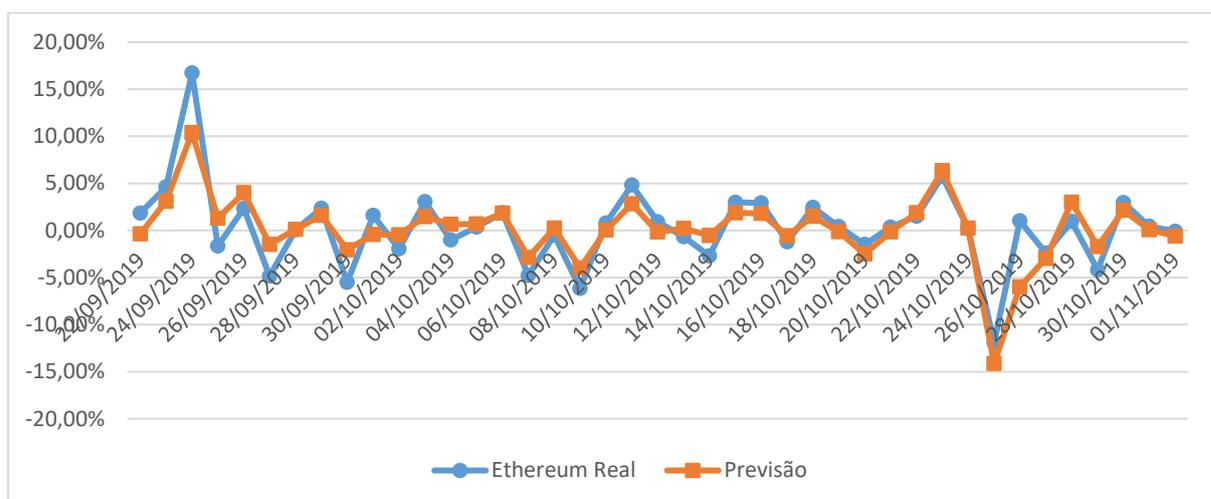
**Fonte: Autoria própria**

Os resultados da Tabela 10 acima foram significativos, devido aos resultados anteriores da análise de variância, onde não foi rejeitada a hipótese nula em nenhum dos casos e fica evidente que o Ethereum tem a relação mais forte com o Bitcoin, com o menor valor S entre os resultados e o maior % de R2 e R2(aj), mostrando um modelo onde 53,64% explica-se a variação nos dados. A relação de TRON com Bitcoin pelo modelo acima, é mais fraca tendo um S muito maior comparado ao Ethereum e um

R2 muito menor, mostrando um modelo onde pouco se explica a variação de seus dados.

Testando-se a equação da moeda Ethereum mostrada na Tabela 10, comparando-se os dados reais com os previstos pela equação no período do dia 22 de setembro até dia 01 de novembro, foi obtido o seguinte resultado mostrado na Figura 13:

**Figura 13 – Comparação Real e Previsão**



**Fonte: Autoria própria**

**Tabela 11 – Resultado de previsão de movimento de mercado**

Dia	Varição Real	Movimento	Varição Prevista	Movimento	Precisão do Movimento
01/11/2019	-0,09%	Baixa	-0,61%	Baixa	Certo
31/10/2019	0,45%	Alta	0,08%	Alta	Certo
30/10/2019	2,97%	Alta	2,14%	Alta	Certo
29/10/2019	-4,20%	Baixa	-1,71%	Baixa	Certo
28/10/2019	0,94%	Alta	2,98%	Alta	Certo
27/10/2019	-2,40%	Baixa	-2,99%	Baixa	Certo
26/10/2019	1,01%	Alta	-5,99%	Baixa	<b>Errado</b>
25/10/2019	-11,92%	Baixa	-14,12%	Baixa	Certo
24/10/2019	0,21%	Alta	0,25%	Alta	Certo

**Fonte: Autoria própria**

A Tabela 11 mostra a precisão no acerto do movimento do mercado, foi obtido um resultado de 75,6% de acerto, onde dos 41 dias testados, o modelo acertou 31 vezes o movimento de alta e baixa.

## 5. CONCLUSÃO

Após finalizadas as análises estatísticas e feitas as devidas análises com os resultados obtidos concluímos que de acordo com o que foi proposto os objetivos foram alcançados. Foi estabelecido um procedimento visando trabalhar com uma base de dados grande (Big Data), que para a análise desse trabalho ficou definido como a variação percentual do preço no dia de 11 criptomoedas, levando em consideração que as mesmas utilizam o Blockchain para registro de seus dados. Após a definição da base de dados que seria trabalhado foi identificado que os métodos estatísticos mais apropriados seria a correlação e também a regressão linear simples, para a realização destas etapas foi essencial a utilização do software Minitab, que se mostrou uma excelente ferramenta sendo fundamental para que os objetivos fossem alcançados.

Também é importante ressaltar a dificuldade em se ter acesso a base de dados de blockchain em projetos de indústria 4.0 dentro das empresas, por se tratar de dados sigilosos em projetos em implementação e devido a isso a escolha de se usar uma base de criptomoedas para realizar o procedimento. O Blockchain é uma tecnologia ainda em amadurecimento na indústria com um potencial revolucionário enorme que necessita mais estudos e análises no tema. O Big Data é uma realidade nas empresas atuais sendo necessário para profissionais de qualquer área técnica, o conhecimento sobre como tratar uma base de dados e realizar análises mais precisas dentro desse contexto. Diante disso acreditamos que os estudos aqui apresentados são de grande importância e abrem caminhos para posteriores análises e aprofundamento no tema.

## REFERÊNCIAS

ANTONPOULOS, Andreas M. **Mastering Bitcoin Unlocking digital cryptocurrencies**. O'reilly, 2016.

COINMARKETCAP. Disponível em: <<https://coinmarketcap.com/>>. Acesso em: 21 set. 2019.

CULTURA ANALITICA. **Os 5 Vs do Big data** Disponível em: <<https://culturaanalitica.com.br/os-5-vs-big-data/>>. Acesso em: 03 Nov.2019.

FIPECAFI – Fundação Instituto de Pesquisas Contábeis, Atuariais e Financeiras; Luiz J Corrar, Jose Maria Dias Filho (coordenadores). **Análise multivariada: para os cursos de Administração, ciências contábeis e Financeiras** – 1.ed – 2/reimpr – São Paulo: Atlas, 2009.

GERBERT, Philipp et al. **Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries**. 2015. Disponível em: <[https://www.bcg.com/publications/2015/engineered\\_products\\_project\\_business\\_industry\\_4\\_future\\_productivity\\_growth\\_manufacturing\\_industries.aspx](https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx)>. Acesso em: 30 out. 2018.

GOMES, Jefferson O. et al. Desafios para indústria 4.0 no Brasil. **Confederação Nacional da Indústria**, Brasília, 2016.

GREENBERG, Andy. Crypto currency, 20 abr 2011. Disponível em: <<https://www.forbes.com/forbes/2011/0509/technology-psilocybin-bitcoins-gavin-andresen-crypto-currency.html#400e4a35353e>>. Acesso em: 25 out. 2018.

HAIR, Joseph F. Jr. **Análise multivariada de dados**. 5 ed.- Porto Alegre: Bookman 2005.

HEKIMA, **O guia definitivo de Big Data para iniciantes**. 2015. Disponível em: <[www.hekima.com](http://www.hekima.com)>.

IDC, **Extracting Value from Chaos**. June, 2011. Disponível em: <<https://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>>.

KAGERMANN, Henning; WAHLSTER, Wolfgang; HELBIG, Johannes. **Recommendations for implementing the strategic initiative INDUSTRIE 4.0**. April, 2013.

KENDALL, M. **Multivariate analysis**. High Wycombe, Charles Griffin, 1980.

KOLOWICH, Lindsay **What is a Whitepaper [FAQ]** Mar. 2018. Disponível em: <<https://blog.hubspot.com/marketing/what-is-a-whitepaper-examples-for-business>>. Acesso em: 25. Out 2018.

LWSISTEMAS. **LTW Sistemas – Conheça os 10 pilares da indústria 4.0** Disponível em: <<https://www.lwtsistemas.com.br/10-pilares-da-industria-4-0/>>. Acesso em: 03 Nov.2019.

MARQUESONE, Rosangela F. P. **Big Data: Técnicas e tecnologias para extração de valor dos dados**. 1. ed. Casa do Código, 2016.

MORAIS, C. Mesquita. **Escalas de Medida, Estatística Descritiva e Inferência Estatística-** Bragança: 2005.

MURRAY, N. Rothbard. **O QUE O GOVERNO FEZ COM O NOSSO DINHEIRO?** São Paulo: Instituto Ludwig Von Mises Brasil 2013. P. 93

NAKAMOTO, Satoshi. **Bitcoin: A Peer-to-Peer Electronic Cash System**. 2009.

NEGRI, Elisa; FUMAGALLI, Luca; MACCHI, Marco. A review of the roles of Digital Twin in CPS-based production systems. **Procedia Manufacturing**. Modena, Itália. v.11, n. June, p. 939-948, 2017.

O'BYRNE, Ian W. **What is blockchain?** 2016 Disponível em: <<https://medium.com/badge-chain/what-is-lockchain-5e4498f05c20>>. Acesso em: 25 out. 2018.

REIS, Elizabeth. **Estatística descritiva**. Lisboa: Sílabo, ed. 7, 2008.

REIS, Edna. **Tabela F**. Disponível em: <<http://www.est.ufmg.br/~edna/Tabela-F.pdf>> Acesso em: 03 Nov.2019.

SANTOS, Cleiton F. et al. Mapping the conceptual relationship among data analysis, knowledge generation and decision-making in industrial processes. **Procedia Manufacturing**. Modena, Itália. v.11, n. June, p. 939-948, 2017.

SCHWAB, Klaus. **A quarta revolução industrial**. São Paulo: Edipro, 2016.

SHIMAKURA, Silvia. E. **Interpretação do coeficiente de correlação**. Ago.2006  
Disponível em: <<http://leg.ufpr.br/~silvia/CE003/node74.html>>.

STANFORD.EDU **Linear Regression** Disponível em:  
<[https://lagunita.stanford.edu/c4x/HumanitiesScience/StatLearning/asset/linear\\_regression.pdf](https://lagunita.stanford.edu/c4x/HumanitiesScience/StatLearning/asset/linear_regression.pdf)>. Acesso em: 01 nov. 2019

ULRICH, Fernando. **Bitcoin: a moeda na era digital**. São Paulo: Instituto Ludwig von Misses Brasil, 2014.

VIALI, Lori. **Série estatística multivariada – texto I**. Rio grande do sul. PucRS

WHITAKER, Amy. **The Eureka Moment That Made Bitcoin Possible**. May.2018.

## APÊNDICE A – BASE DE DADOS

Date	Bitcoin	Ethere	XRP	Bitcoin Ca	Litecoir	EOS	Binance Co	Stellar	Cardan	Moner	TRON
Sep 21, 2019	1,61%	1,09%	1,70%	1,03%	2,21%	0,00%	1,40%	3,76%	0,79%	1,58%	-4,78%
Sep 20, 2019	0,82%	1,37%	2,64%	2,61%	2,35%	-0,75%	1,20%	9,23%	0,29%	2,73%	1,41%
Sep 19, 2019	-0,65%	-4,75%	3,60%	0,29%	1,10%	2,69%	2,83%	2,10%	1,49%	7,01%	1,98%
Sep 18, 2019	0,48%	-1,61%	-10,42%	-1,67%	-5,04%	-0,25%	-5,95%	-30,59%	-8,87%	-10,20%	-7,31%
Sep 17, 2019	0,39%	-5,83%	-9,26%	-3,89%	-2,25%	0,00%	-2,99%	-8,71%	-4,03%	1,41%	-6,98%
Sep 16, 2019	0,68%	-3,91%	0,24%	-1,07%	-3,20%	-0,25%	0,97%	-0,77%	-0,55%	0,46%	-0,63%
Sep 15, 2019	0,08%	-0,91%	-0,18%	0,12%	0,38%	-1,50%	2,23%	-0,63%	-0,63%	0,50%	1,35%
Sep 14, 2019	-0,12%	-3,74%	-2,22%	-2,10%	-2,23%	-7,51%	-0,81%	-1,53%	-1,67%	1,03%	-1,80%
Sep 13, 2019	0,53%	-0,07%	-0,30%	0,73%	0,35%	0,00%	0,19%	-2,45%	-0,64%	-2,33%	-4,38%
Sep 12, 2019	-2,29%	-1,24%	0,45%	-0,27%	1,04%	1,06%	-0,14%	2,31%	-1,02%	-0,99%	0,62%
Sep 11, 2019	-0,55%	0,62%	1,06%	1,46%	1,35%	0,53%	4,61%	-2,39%	3,24%	-2,86%	4,28%
Sep 10, 2019	2,13%	0,78%	0,39%	0,33%	-0,98%	2,58%	1,70%	2,42%	0,84%	3,91%	0,76%
Sep 09, 2019	1,04%	0,12%	1,40%	0,77%	0,61%	-2,37%	1,24%	1,72%	-0,48%	2,91%	0,62%
Sep 08, 2019	0,73%	-1,73%	-0,88%	-1,89%	-2,22%	-5,88%	0,04%	-2,17%	-1,10%	1,35%	-2,62%
Sep 07, 2019	-1,58%	-4,88%	-3,28%	-4,66%	-5,82%	-10,53%	-1,85%	-1,38%	-3,99%	-3,66%	-8,22%
Sep 06, 2019	2,13%	2,42%	1,53%	1,87%	0,24%	1,52%	2,68%	3,01%	0,05%	0,49%	5,00%
Sep 05, 2019	0,12%	0,99%	1,37%	0,82%	2,84%	0,91%	-4,31%	2,78%	2,34%	-2,72%	3,73%
Sep 04, 2019	0,25%	1,95%	1,25%	1,89%	2,73%	1,49%	3,24%	1,65%	2,71%	1,31%	2,31%
Sep 03, 2019	-2,69%	-0,64%	-0,88%	-2,33%	-3,19%	-1,20%	1,10%	0,07%	-3,07%	-2,35%	-0,09%
Sep 02, 2019	-6,04%	-3,87%	-1,07%	-4,16%	-1,26%	-1,22%	-5,56%	-1,22%	-1,81%	-1,99%	-2,02%
Sep 01, 2019	-1,32%	0,48%	0,43%	-0,84%	-2,45%	1,50%	-1,46%	0,74%	0,91%	-6,32%	-0,35%
Aug 31, 2019	-0,35%	-2,12%	-1,17%	-0,10%	-0,34%	-3,42%	4,23%	-1,01%	-0,37%	0,47%	0,31%
Aug 30, 2019	-0,88%	0,39%	0,29%	0,32%	-0,39%	0,00%	-1,51%	1,20%	-0,24%	-1,01%	-0,44%
Aug 29, 2019	2,53%	2,55%	0,16%	3,94%	4,84%	2,13%	7,60%	5,20%	2,48%	7,15%	2,92%
Aug 28, 2019	4,40%	7,26%	4,50%	5,14%	7,50%	7,58%	7,78%	3,55%	8,10%	7,96%	9,18%
Aug 27, 2019	1,81%	0,76%	0,47%	1,00%	1,47%	0,84%	1,65%	2,03%	-0,27%	1,91%	0,61%
Aug 26, 2019	-2,41%	-1,17%	-0,31%	-1,90%	-2,13%	-0,56%	0,11%	1,90%	-0,49%	0,91%	-0,88%
Aug 25, 2019	0,22%	2,30%	0,93%	0,65%	1,58%	3,00%	1,36%	-6,58%	1,46%	0,47%	1,25%
Aug 24, 2019	2,38%	1,74%	1,44%	2,69%	2,59%	0,27%	2,43%	2,47%	-0,75%	1,69%	1,57%
Aug 23, 2019	-2,68%	-1,88%	-2,21%	-1,77%	-2,47%	-1,10%	-0,30%	-4,19%	-0,77%	0,10%	-3,33%
Aug 22, 2019	0,11%	-2,35%	-1,93%	-2,72%	-0,89%	-4,00%	-0,37%	2,24%	-4,79%	-2,11%	-5,07%
Aug 21, 2019	5,82%	4,95%	3,45%	4,20%	3,03%	4,89%	4,19%	3,72%	4,29%	6,06%	5,18%
Aug 20, 2019	1,40%	3,08%	2,60%	2,83%	2,72%	1,60%	2,22%	2,27%	2,10%	4,89%	2,65%
Aug 19, 2019	-5,47%	-4,38%	-0,37%	-2,80%	-1,21%	-0,81%	-3,08%	-1,42%	0,84%	-2,36%	-1,41%
Aug 18, 2019	-1,10%	-4,65%	-6,37%	-3,00%	-4,71%	-3,93%	-1,60%	-4,21%	-6,97%	-7,34%	-5,25%
Aug 17, 2019	1,23%	-0,09%	-1,46%	1,06%	2,82%	0,84%	0,61%	1,52%	-1,12%	0,69%	0,19%
Aug 16, 2019	-0,53%	1,70%	1,17%	1,38%	1,83%	1,10%	1,50%	1,93%	1,46%	0,02%	3,64%
Aug 15, 2019	-2,72%	-0,97%	-0,45%	-0,77%	-0,34%	1,09%	-3,47%	-1,54%	-0,91%	-5,03%	0,63%
Aug 14, 2019	7,69%	10,54%	11,25%	9,32%	9,90%	9,61%	7,98%	6,73%	10,57%	8,22%	12,89%
Aug 13, 2019	4,30%	1,24%	1,34%	-4,07%	1,56%	1,45%	2,32%	1,63%	0,70%	6,22%	1,71%
Aug 12, 2019	1,26%	2,21%	0,61%	1,76%	4,40%	1,43%	-0,30%	3,12%	3,39%	1,19%	0,17%
Aug 11, 2019	-1,53%	-4,53%	-1,12%	-6,43%	-5,03%	-2,95%	-1,28%	-5,55%	-5,75%	-1,54%	-1,25%
Aug 10, 2019	4,28%	1,80%	-0,50%	0,22%	-1,22%	-4,09%	0,87%	-2,10%	-9,11%	2,63%	-3,47%
Aug 09, 2019	0,76%	4,72%	3,32%	5,11%	6,27%	6,01%	2,95%	6,04%	8,22%	1,92%	9,60%
Aug 08, 2019	-0,10%	2,46%	1,32%	1,74%	0,97%	1,65%	-4,97%	2,26%	-0,12%	2,36%	-0,97%
Aug 07, 2019	-4,06%	-0,20%	-0,15%	-1,74%	2,29%	-0,95%	-6,19%	-0,54%	2,75%	-8,64%	-0,02%
Aug 06, 2019	2,82%	3,51%	3,92%	3,81%	3,94%	6,68%	1,43%	5,41%	5,80%	4,14%	6,11%
Aug 05, 2019	-7,71%	-5,20%	-1,20%	-3,92%	-4,02%	-5,90%	-1,71%	-0,63%	-0,73%	-6,20%	-3,03%
Aug 04, 2019	-1,37%	-0,04%	-0,94%	-0,18%	1,48%	1,40%	1,15%	0,33%	2,59%	0,01%	-2,27%
Aug 03, 2019	-2,88%	-2,11%	-1,36%	-1,87%	0,46%	-2,63%	0,68%	-0,51%	-0,60%	-5,08%	-0,78%
Aug 02, 2019	-1,12%	0,00%	1,70%	0,58%	4,29%	3,23%	2,57%	1,29%	2,47%	-2,45%	2,26%
Aug 01, 2019	-3,20%	0,34%	0,88%	-0,66%	-0,62%	2,26%	-4,16%	1,14%	2,14%	-1,69%	1,01%

Date	Bitcoin	Ethereum	XRP	Bitcoin Cash	Litecoin	EOS	Binance Coin	Stellar	Cardano	Monero	TRON
Nov 12, 2017	5,49%	2,15%	6,12%	-3,11%	5,30%	6,03%	8,98%	14,51%	12,75%	-3,56%	11,12%
Nov 11, 2017	3,94%	-5,39%	-2,18%	-33,83%	-4,97%	-12,62%	6,70%	0,09%	-2,03%	-13,12%	1,14%
Nov 10, 2017	7,74%	6,68%	5,39%	-54,56%	8,04%	12,71%	10,45%	16,97%	15,83%	12,98%	13,79%
Nov 09, 2017	4,07%	-3,97%	0,19%	-5,14%	-2,85%	-4,42%	2,93%	-18,21%	-27,40%	-7,33%	-1,74%
Nov 08, 2017	-4,46%	-5,03%	-3,21%	-1,17%	-2,39%	-18,82%	-12,78%	-10,13%	-14,94%	-13,52%	-7,59%
Nov 07, 2017	-1,73%	1,31%	-2,56%	-2,26%	-11,47%	-3,29%	-15,29%	-6,32%	-1,60%	4,29%	-1,82%
Nov 06, 2017	5,14%	-0,83%	-1,99%	4,39%	-0,75%	14,30%	-2,61%	-8,19%	-4,31%	-19,09%	3,98%
Nov 05, 2017	-0,04%	1,26%	0,60%	-1,74%	0,51%	1,80%	7,32%	1,76%	6,43%	1,10%	-4,37%
Nov 04, 2017	-3,01%	1,64%	2,08%	2,35%	2,10%	6,67%	-12,33%	1,92%	2,73%	0,68%	4,49%
Nov 03, 2017	-1,70%	-5,97%	-0,93%	-6,47%	-2,93%	-8,04%	-13,18%	-1,02%	-5,65%	-5,34%	3,34%
Nov 02, 2017	-4,44%	1,14%	-5,79%	-11,28%	-2,20%	-6,67%	-3,17%	-2,15%	7,68%	2,22%	8,40%
Nov 01, 2017	-5,07%	4,60%	3,06%	-21,34%	4,59%	-33,61%	5,97%	6,92%	21,08%	2,20%	9,78%
Oct 31, 2017	-5,49%	0,49%	0,91%	2,30%	1,10%	-9,69%	2,24%	7,67%	-6,62%	1,25%	-1,96%
Oct 30, 2017	-0,26%	-0,97%	0,18%	2,23%	0,79%	-6,56%	-6,50%	-3,49%	-0,94%	-0,28%	5,66%
Oct 29, 2017	-6,94%	-2,94%	-0,71%	-7,48%	-4,46%	-24,14%	-5,08%	2,49%	-10,75%	-2,76%	1,07%
Oct 28, 2017	0,60%	0,54%	0,37%	-14,56%	1,25%	-2,70%	2,54%	9,47%	1,80%	0,69%	1,96%
Oct 27, 2017	2,01%	-0,31%	0,43%	-8,65%	0,61%	-0,60%	3,31%	1,28%	2,79%	2,07%	-7,43%
Oct 26, 2017	-2,73%	0,50%	0,40%	-1,83%	1,29%	-2,90%	3,97%	1,99%	1,01%	-0,47%	6,21%
Oct 25, 2017	-4,09%	0,17%	1,21%	-1,70%	-0,97%	-2,10%	-0,81%	3,92%	-1,81%	0,41%	-14,19%
Oct 24, 2017	6,89%	-3,90%	-5,66%	-3,48%	-1,82%	-4,49%	-5,93%	-6,33%	-7,32%	-3,80%	-8,54%
Oct 23, 2017	1,26%	2,76%	3,23%	4,56%	3,26%	4,72%	4,07%	-8,42%	8,06%	1,47%	5,71%
Oct 22, 2017	0,47%	1,70%	3,21%	-2,52%	2,51%	1,62%	3,97%	14,30%	2,93%	1,81%	-12,01%
Oct 21, 2017	-0,58%	1,11%	-0,39%	1,04%	3,48%	2,49%	-0,81%	-33,23%	7,69%	2,53%	-17,67%
Oct 20, 2017	-5,31%	1,31%	2,11%	1,05%	-0,87%	4,58%	3,13%	14,37%	-13,52%	-2,06%	-1,36%
Oct 19, 2017	-2,23%	2,05%	1,41%	2,49%	1,78%	-3,86%	3,76%	11,32%	-1,23%	1,16%	8,55%
Oct 18, 2017	0,23%	0,83%	5,49%	9,23%	-1,91%	1,49%	4,32%	23,31%	1,94%	1,81%	-7,82%
Oct 17, 2017	2,37%	4,92%	9,38%	-18,50%	8,03%	1,82%	3,47%	-25,56%	6,00%	4,85%	9,83%
Oct 16, 2017	-0,67%	0,99%	2,72%	0,51%	1,67%	1,75%	7,69%	-94,75%	3,31%	-0,96%	5,73%
Oct 15, 2017	2,70%	0,86%	-2,50%	2,32%	-2,09%	1,98%	-4,00%	-9,45%	8,59%	5,07%	3,87%
Oct 14, 2017	-3,34%	-0,23%	1,47%	-0,13%	-7,64%	0,34%	-4,76%	2,75%	4,17%	-5,44%	-13,25%
Oct 13, 2017	-3,35%	-11,40%	-4,92%	-2,33%	1,85%	1,29%	-7,46%	-1,76%	-29,09%	-9,06%	9,89%
Oct 12, 2017	-12,78%	-0,21%	6,28%	0,72%	-18,12%	3,04%	6,29%	8,79%	-16,48%	0,54%	5,10%
Oct 11, 2017	-0,78%	-0,76%	-1,45%	1,66%	-0,28%	-9,66%	-2,14%	-3,07%	-4,84%	-0,78%	-7,91%
Oct 10, 2017	-0,12%	-0,76%	-3,13%	-2,67%	-1,14%	0,42%	2,80%	-8,68%	2,63%	-1,41%	-1,78%
Oct 09, 2017	-3,41%	3,65%	9,66%	8,34%	5,50%	7,69%	-15,70%	6,38%	-8,70%	4,74%	15,04%
Oct 08, 2017	-4,08%	0,79%	-17,47%	4,97%	-1,24%	-2,91%	22,78%	10,12%	2,16%	1,77%	1,36%
Oct 07, 2017	-1,32%	-0,72%	-1,99%	0,78%	-0,84%	1,88%	1,86%	-13,64%	-13,72%	-0,84%	17,90%
Oct 06, 2017	-1,07%	-4,55%	1,55%	-2,13%	-0,93%	1,13%	5,95%	-25,56%	13,20%	0,78%	9,79%
Oct 05, 2017	-2,33%	-1,06%	-9,94%	0,74%	-0,53%	-1,14%	-3,07%	-22,68%	2,10%	-1,51%	-1,56%
Oct 04, 2017	2,08%	0,03%	-5,69%	11,32%	1,68%	4,82%	13,51%	4,96%	-5,11%	2,73%	-31,63%
Oct 03, 2017	2,06%	1,69%	-0,05%	4,18%	1,95%	3,31%	7,61%	4,19%	19,18%	-1,57%	10,95%
Oct 02, 2017	-0,31%	1,65%	1,65%	-1,28%	2,14%	7,49%	-28,10%	5,37%	-5,38%	2,16%	-51,68%
Oct 01, 2017	-1,44%	-0,26%	-4,51%	4,21%	1,09%	1,94%	-19,53%	0,90%	-0,49%	2,48%	18,27%
Sep 30, 2017	-4,14%	-3,42%	-1,91%	0,92%	-4,77%	-0,41%	-0,79%	-5,29%	-0,49%	-1,40%	-14,02%
Sep 29, 2017	0,20%	2,56%	1,27%	2,43%	2,75%	12,20%	7,30%	4,43%	-0,49%	2,88%	31,64%
Sep 28, 2017	0,53%	2,39%	3,16%	1,95%	3,77%	-28,87%	-20,00%	3,16%	-0,49%	3,70%	-13,38%
Sep 27, 2017	-7,90%	-6,78%	-9,72%	-3,50%	-9,11%	-13,00%	-33,22%	-16,90%	-0,49%	-8,94%	-47,52%
Sep 26, 2017	0,92%	1,68%	-2,28%	0,79%	0,27%	1,73%	6,33%	-2,58%	-0,49%	0,72%	2,19%
Sep 25, 2017	-6,64%	-3,58%	-3,40%	-6,23%	-8,32%	-4,44%	-16,94%	-4,52%	-0,49%	-4,78%	21,06%
Sep 24, 2017	2,98%	1,30%	1,34%	1,81%	3,31%	5,85%	-0,96%	2,95%	-0,49%	2,46%	4,26%
Sep 23, 2017	-4,48%	-8,15%	-2,60%	-3,18%	-2,30%	-1,03%	-7,50%	-10,61%	-0,49%	-4,34%	-21,30%
Sep 22, 2017	-0,07%	-2,40%	-1,13%	0,68%	-3,09%	-4,04%	-2,44%	-0,72%	-0,49%	-2,32%	6,16%

## ANEXO A – TABELA F

Tabela 5. Limites unilaterais da distribuição F de Fisher-Snedecor ao nível de 5% de probabilidade.

GL V2	V1																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	40	60	120	240
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.0	243.9	244.7	245.4	245.9	248.0	251.1	252.2	253.3	253.8
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.385	19.396	19.405	19.412	19.419	19.424	19.429	19.446	19.471	19.479	19.487	19.492
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.785	8.763	8.745	8.729	8.715	8.703	8.680	8.594	8.572	8.549	8.538
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.936	5.912	5.891	5.873	5.858	5.803	5.717	5.688	5.658	5.643
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.704	4.678	4.655	4.636	4.619	4.558	4.464	4.431	4.398	4.382
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	4.027	4.000	3.976	3.956	3.938	3.874	3.774	3.740	3.705	3.687
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.603	3.575	3.550	3.529	3.511	3.445	3.340	3.304	3.267	3.249
8	5.318	4.459	4.066	3.838	3.688	3.581	3.500	3.438	3.388	3.347	3.313	3.284	3.259	3.237	3.218	3.150	3.043	3.005	2.967	2.947
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.102	3.073	3.048	3.025	3.006	2.936	2.826	2.787	2.748	2.727
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.943	2.913	2.887	2.865	2.845	2.774	2.661	2.621	2.580	2.559
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.818	2.788	2.761	2.739	2.719	2.646	2.531	2.490	2.448	2.426
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.717	2.687	2.660	2.637	2.617	2.544	2.426	2.384	2.341	2.319
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.635	2.604	2.577	2.554	2.533	2.459	2.339	2.297	2.252	2.230
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.565	2.534	2.507	2.484	2.463	2.388	2.266	2.223	2.178	2.155
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.507	2.475	2.448	2.424	2.403	2.328	2.204	2.160	2.114	2.090
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494	2.456	2.425	2.397	2.373	2.352	2.276	2.151	2.106	2.059	2.035
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.413	2.381	2.353	2.329	2.308	2.230	2.104	2.058	2.011	1.986
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.374	2.342	2.314	2.290	2.269	2.191	2.063	2.017	1.968	1.943
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.340	2.308	2.280	2.256	2.234	2.155	2.026	1.980	1.930	1.905
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.310	2.278	2.250	2.225	2.203	2.124	1.994	1.946	1.896	1.870
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321	2.283	2.250	2.222	2.197	2.176	2.096	1.965	1.916	1.866	1.839
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297	2.259	2.226	2.198	2.173	2.151	2.071	1.938	1.889	1.838	1.811
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275	2.236	2.204	2.175	2.150	2.128	2.048	1.914	1.865	1.813	1.785
24	4.260	3.403	3.009	2.776	2.620	2.508	2.423	2.355	2.300	2.255	2.216	2.183	2.155	2.130	2.108	2.027	1.892	1.842	1.790	1.762
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236	2.198	2.165	2.136	2.111	2.089	2.007	1.872	1.822	1.768	1.740
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220	2.181	2.148	2.119	2.094	2.072	1.990	1.853	1.803	1.749	1.720
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204	2.166	2.132	2.103	2.078	2.056	1.974	1.836	1.785	1.731	1.702
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190	2.151	2.118	2.089	2.064	2.041	1.959	1.820	1.769	1.714	1.685
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177	2.138	2.104	2.075	2.050	2.027	1.945	1.806	1.754	1.698	1.669
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.126	2.092	2.063	2.037	2.015	1.932	1.792	1.740	1.683	1.654
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	2.038	2.003	1.974	1.948	1.924	1.839	1.693	1.637	1.577	1.544
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026	1.986	1.952	1.921	1.895	1.871	1.784	1.634	1.576	1.511	1.476
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.952	1.917	1.887	1.860	1.836	1.748	1.594	1.534	1.467	1.430
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951	1.910	1.875	1.845	1.817	1.793	1.703	1.545	1.482	1.411	1.370
100	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927	1.886	1.850	1.819	1.792	1.768	1.676	1.515	1.450	1.376	1.333
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910	1.869	1.834	1.803	1.775	1.750	1.659	1.495	1.429	1.352	1.307
240	3.881	3.033	2.642	2.409	2.252	2.136	2.048	1.977	1.919	1.870	1.829	1.793	1.761	1.733	1.708	1.614	1.445	1.375	1.290	1.237