

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
NÚCLEO DE EDUCAÇÃO A DISTÂNCIA
Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data

Roberta Fonseca Santos de Paula

**PREVISÃO DE PREÇOS DE AÇÕES COM SÉRIES TEMPORAIS DA BOLSA DE
VALORES**

Belo Horizonte
2022

Roberta Fonseca Santos de Paula

**PREVISÃO DE PREÇOS DE AÇÕES COM SÉRIES TEMPORAIS DA BOLSA DE
VALORES**

Trabalho de Conclusão de Curso apresentado
ao Curso de Especialização em Ciência de
Dados e Big Data como requisito parcial à
obtenção do título de especialista.

Belo Horizonte

2022

SUMÁRIO

1. Introdução	4
1.1. Contextualização	4
1.2. O problema proposto	4
2. Coleta de Dados	5
3. Processamento/Tratamento de Dados	6
4. Análise e Exploração dos Dados	7
4.1. Análise Exploratória da BBAS3.SA.....	7
4.2. Análise Exploratória da BBDC4.SA	12
4.3. Análise Exploratória da ITUB4.SA	18
4.4. Análise das Ações em Conjunto	22
5. Criação de Modelos de Machine Learning	24
5.1. Modelo Preditivo Com Facebook Prophet	25
5.2. Modelo Preditivo Com Auto-Arima	26
5.3. Modelo Preditivo Com RNN (LSTM – Long Short Term Memory)	30
6. Apresentação dos Resultados	31
6.1. Modelo Preditivo Com Facebook Prophet	31
6.2. Modelo Preditivo Com Auto-Arima	35
6.3. Modelo Preditivo Com RNN (LSTM – Long Short Term Memory)	37
7. Links	38
REFERÊNCIAS.....	39

1. Introdução

1.1. Contextualização

O Ibovespa é o principal indicador de desempenho das ações negociadas na B3 e reúne as empresas mais importantes do mercado de capitais brasileiro. Foi criado em 1968 e, ao longo desse tempo, consolidou-se como referência para investidores ao redor do mundo.

As movimentações da bolsa de valores são de suma importância para os investidores, influenciando sua tomada de decisão ao realizar investimentos.

No entanto, são dados que, essencialmente, são muito dinâmicos e caóticos por natureza, visto que os preços das ações são afetados por diversos fatores, como preço da moeda, política, relações comerciais e internacionais, índices econômicos, entre outros.

O mercado financeiro caracteriza-se por incertezas, que são associadas com as expectativas de curto, médio e longo prazos.

Com isso, é difícil prever a oscilação das ações ao longo do tempo. Mas é possível utilizar modelos preditivos que podem capturar as dinâmicas do mercado, com o objetivo de diminuir as incertezas e observar as tendências dos preços das ações para melhor tomada de decisão por parte dos investidores.

Para isso, foi desenvolvido um script em Python utilizando diversas bibliotecas, como Yahoo Finance API, Keras, Facebook Prophet, Arima, além de bibliotecas de uso comum da linguagem.

1.2. O problema proposto

Este estudo consiste no uso de análise exploratória e modelagem preditiva para extração de informações importantes das séries temporais do preço de fechamento das ações para auxiliar os investidores a obterem maior rentabilidade no mercado de ações.

[Digite aqui]

O trabalho tem como objetivo criar modelos preditivos para a ação do Banco do Brasil (BBAS3.SA) em um intervalo de um ano, utilizando séries temporais da bolsa de valores, obtidas através da API do Yahoo Finance.

Para isso, serão analisadas as séries temporais das ações do Banco do Brasil (BBAS3.SA), Banco Bradesco (BBDC4.SA) e Banco Itaú (ITUB4.SA) para treinar os modelos de predição.

Para facilitar o entendimento do problema e da solução, será utilizada a técnica dos 5-Ws, que consiste em responder as seguintes perguntas:

Why? A bolsa de valores é muito relevante para a economia do país, e entender suas movimentações e oscilações é muito importante para auxiliar os investidores a terem insights para tomada de decisões e, com isso, obterem maior rentabilidade no Mercado de Ações.

Who? Os dados analisados são de três bancos importantes no Brasil (Banco do Brasil, Banco Bradesco e Banco Itaú), extraídos da API do Yahoo Finance.

What? Prever o preço de fechamento da ação do Banco do Brasil em um intervalo de um ano utilizando séries temporais da bolsa de valores, obtidas através da API do Yahoo Finance.

Where? Os dados analisados são nacionais. Foram extraídos da API Yahoo Finance, que não é nacional, mas os dados pertencem a empresas brasileiras.

When? Os dados analisados são do período de 01/01/2015 a 31/12/2019, totalizando 4 anos.

2. Coleta de Dados

Para o desenvolvimento desse trabalho, foi utilizado o Jupyter Notebook, e foi utilizada a API Yahoo Finance para obter os datasets das ações do Banco do Brasil (BBAS3.SA), Banco Bradesco (BBDC4.SA) e Banco Itaú (ITUB4.SA) da bolsa de valores B3. Para a leitura dos dados, foi utilizada a biblioteca Pandas Datareader em comunicação com a API Yahoo Finance.

[Digite aqui]

Os dados são obtidos ao executar o script, bastando apenas passar como parâmetros o nome da ação, data de início e data de fim

O dataset obtido para cada ação tem o formato descrito na tabela abaixo:

Nome da coluna/campo	Descrição	Tipo
Date	Data da negociação	Pandas DatetimeIndex
High	Maior preço da ação no dia	Pandas Series
Low	Menor preço da ação no dia	Pandas Series
Open	Preço de abertura da ação no dia	Pandas Series
Close	Preço de fechamento da ação no dia	Pandas Series
Volume	Volume total negociado no dia	Pandas Series
Adj Close	Valor ajustado da ação no dia. Este valor considera desconto de dividendos, splits, inplits, etc.	Pandas Series

Neste trabalho, foi utilizado o preço de fechamento (Close) nos modelos de predição.

3. Processamento/Tratamento de Dados

Após o processo de leitura dos dados, obtém-se os seguintes datasets:

Nome do Dataset	Descrição
<code>train_dataset_bbas3</code>	Conjunto de dados de treinamento da ação BBAS3 no período de 01/01/2015 a 31/12/2019.
<code>train_dataset_bbd4</code>	Conjunto de dados de treinamento da ação BBDC4 no período de 01/01/2015 a 31/12/2019.
<code>train_dataset_itub4</code>	Conjunto de dados de treinamento da ação ITUB4 no período de 01/01/2015 a 31/12/2019.

[Digite aqui]

test_dataset_bbas3	Conjunto de dados de teste a ser usado nos modelos preditivos da ação BBAS3 no período de 01/01/2020 a 31/12/2020.
--------------------	--

Cada dataset de treinamento possui 1243 registros, e o de teste da BBAS3 possui 248.

Não há registros duplicados, já que cada registro corresponde a uma data diferente. Além disso, os dados correspondem a dias em que a Bolsa de Valores está aberta para negociações, assim, esses registros correspondem aos dias úteis do ano. Também foi verificado que não há nenhum valor nulo nos dados.

4. Análise e Exploração dos Dados

4.1. Análise Exploratória da BBAS3.SA

Após a coleta dos dados, foi realizada uma descrição estatística do dataset, conforme mostra a tabela a seguir:

Cam po	High	Low	Open	Close	Volume	Adj Close
Co unt	1243.0	1243.0	1243.0	1243.0	1243.0	1243.0
Me an	32.170040 22498526	31.210120 695210808	31.703724 85347266	31.681922 76912453	9505695.4 14320193	24.728954 897517653
Std	11.441466 925885765	11.250512 714247257	11.361092 208476675	11.354133 222203913	4955245.6 785919685	10.385393 134496292
Min	12.840000 15258789	12.520000 457763672	12.619999 885559082	12.699999 809265137	0.0	9.1677455 90209961
25 %	23.025000 57220459	22.219999 313354492	22.625	22.590000 15258789	6386600.0	15.706981 182098389

[Digite aqui]

50	30.719999	29.680000	30.299999	30.360000	8633100.0	23.576946
%	313354492	30517578	237060547	610351562		258544922
75	41.795000	40.610000	41.319999	41.210000	11422600.	32.437114
%	076293945	61035156	69482422	99182129	0	71557617
Ma	55.909999	54.819999	55.599998	55.180000	47116600.	46.556953
x	84741211	69482422	474121094	30517578	0	43017578

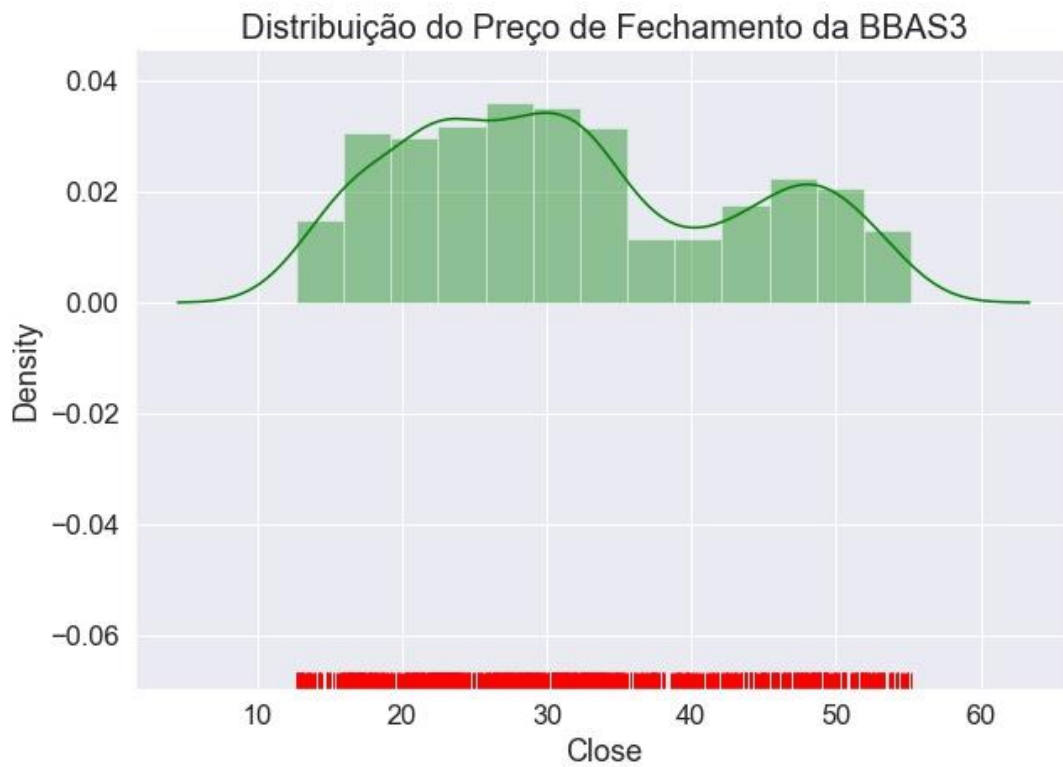
A seguir, verificou-se que não há valores nulos no dataset:

Campo	Valores nulos
High	0
Low	0
Open	0
Close	0
Volume	0
Adj Close	0

Ao plotar o histograma do preço de fechamento da ação, obtém-se o resultado apresentado na figura 1.

[Digite aqui]

Figura 1: Histograma do preço de fechamento da BBAS3



Também foi feita a plotagem da série histórica de preços contemplando o preço máximo, mínimo e o fechamento, conforme mostra a figura 2.

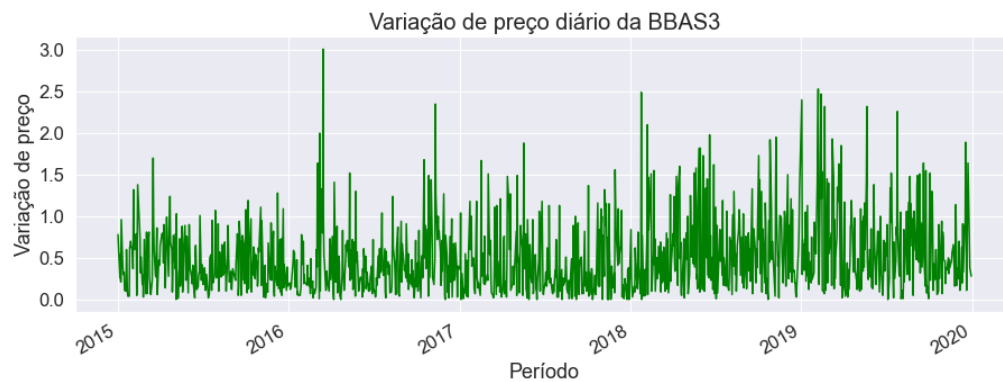
Figura 2: Histórico de preços da BBAS3



Plotando a variação de preço diário (preço de fechamento - preço de abertura), obtém-se o resultado apresentado na figura 3.

[Digite aqui]

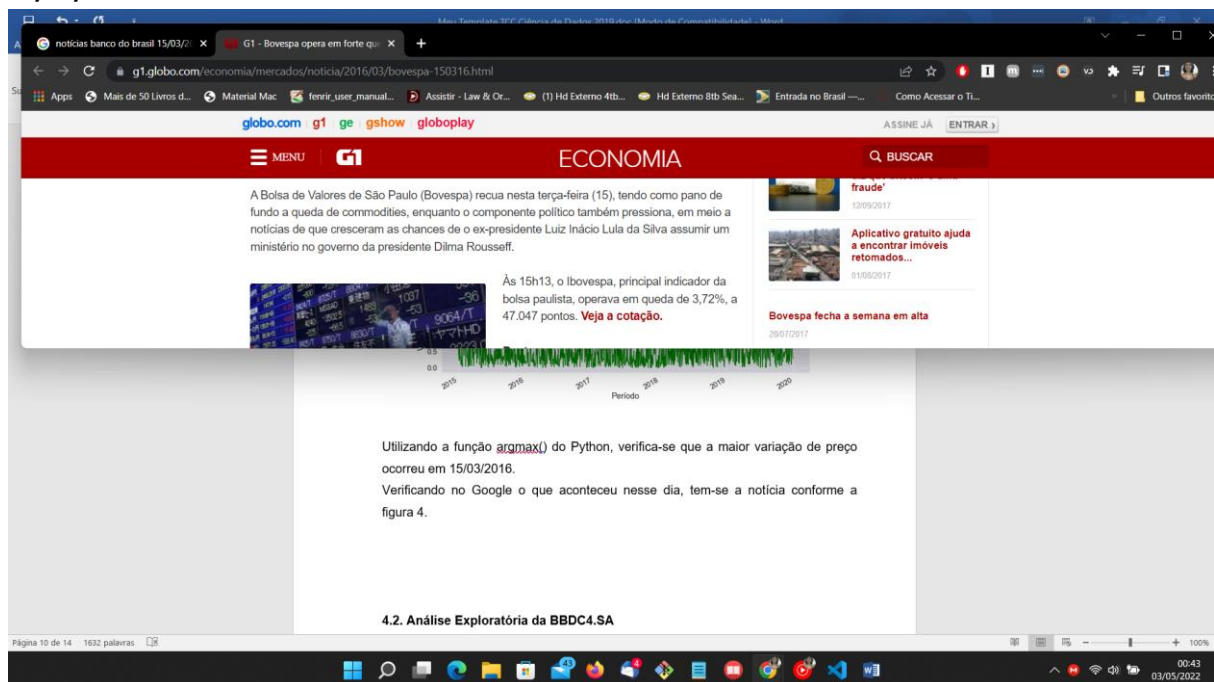
Figura 3: Variação de preço diário (preço de fechamento - preço de abertura) da BBAS3



Utilizando a função `argmax()` do Python, verifica-se que a maior variação de preço ocorreu em 15/03/2016.

Verificando no Google o que aconteceu nesse dia, tem-se a notícia conforme a figura 4.

Figura 4: Notícia da BBAS3 encontrada no Google no dia 15/03/2016



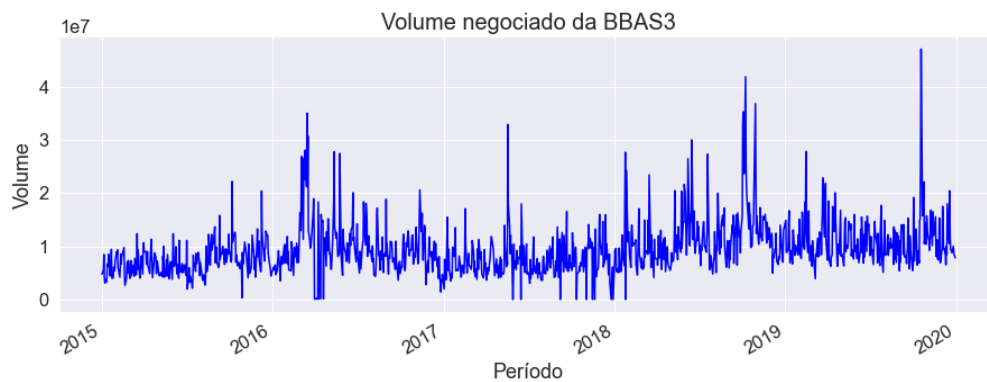
[Notícia da BBAS3 encontrada no Google no dia 15/03/2016](#)

[Digite aqui]

Analisando-se a notícia, percebe-se que o que influenciou essa variação foi a queda de commodities, juntamente com a pressão do ambiente político, já que havia rumores de que o ex-presidente Luiz Inácio Lula da Silva poderia assumir um ministério no governo da então presidente Dilma Rousseff.

Plotando o volume negociado, tem-se o resultado apresentado na figura 5.

Figura 5: Volume negociado da BBAS3

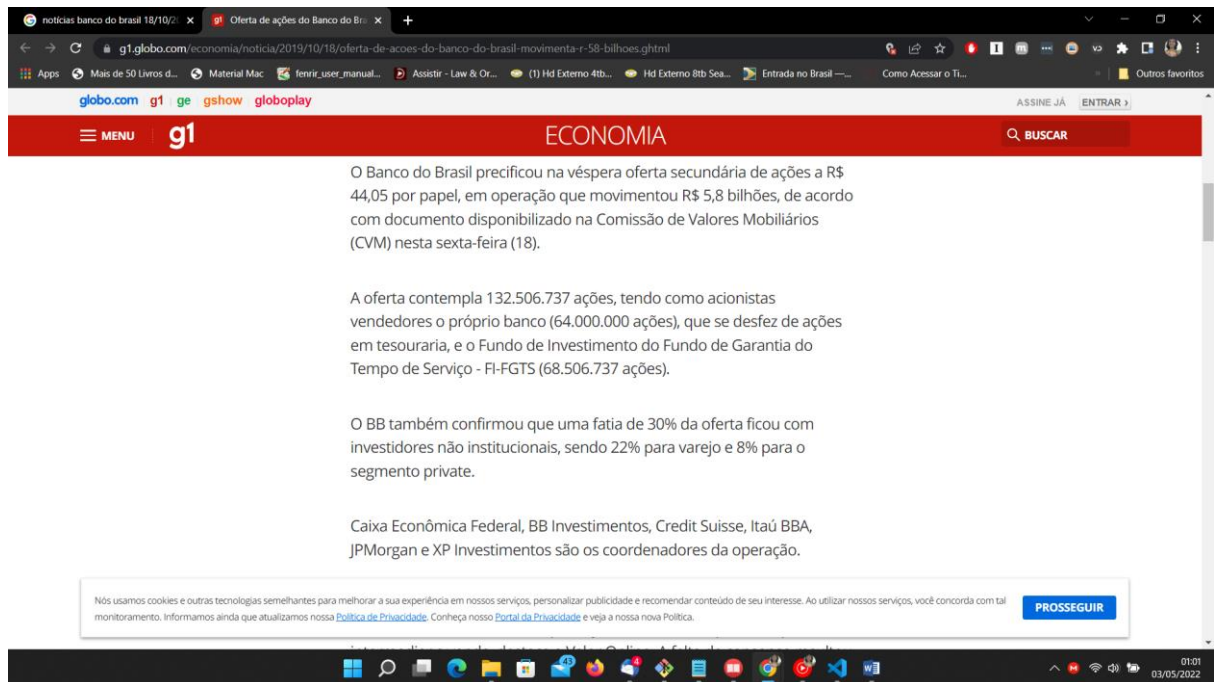


Utilizando a função `argmax()` do Python, verifica-se que o dia com maior volume de negociação foi em 18/10/2019.

Verificando no Google o que aconteceu nesse dia, tem-se a notícia conforme a figura 6.

[Digite aqui]

Figura 6: Notícia da BBAS3 encontrada no Google no dia 18/10/2019



Notícia da BBAS3 encontrada no Google no dia 18/10/2019

Analisando-se a notícia, percebe-se que o que influenciou esse grande volume de negociação foi a grande oferta de ações por parte do Banco do Brasil.

4.2. Análise Exploratória da BBDC4.SA

Após a coleta dos dados, foi realizada uma descrição estatística do dataset, conforme mostra a tabela a seguir:

Ca mp o	High	Low	Open	Close	Volume	Adj Close
Co unt	1243.0	1243.0	1243.0	1243.0	1243.0	1243.0
Me an	18.325319 38606563	17.854553 353853603	18.101317 034358473	18.103179 69840089	22002657. 414320193	14.731994 39067948

[Digite aqui]

Std	5.4749030 48081267	5.3766198 01369015	5.4356497 474707535	5.4337839 35198473	10723897. 665898316	5.2422113 32704079
Min	8.1472396 85058594	7.9637861 251831055	8.0061206 81762695	8.0014171 6003418	0.0	6.0822830 20019531
25 %	14.191736 221313477	13.702917 575836182	13.935449 600219727	13.910910 12954712	15023448. 5	10.263040 06576538
50 %	17.173648 834228516	16.750904 083251953	17.001342 7734375	16.956325 53100586	20218811. 0	13.646457 67211914
75 %	22.357307 43408203	21.779875 75531006	22.163785 934448242	22.089793 20526123	26613478. 0	18.389400 482177734
Ma x	30.007513 04626465	29.316303 253173828	29.752065 658569336	29.609315 872192383	106029201 .0	25.589248 657226562

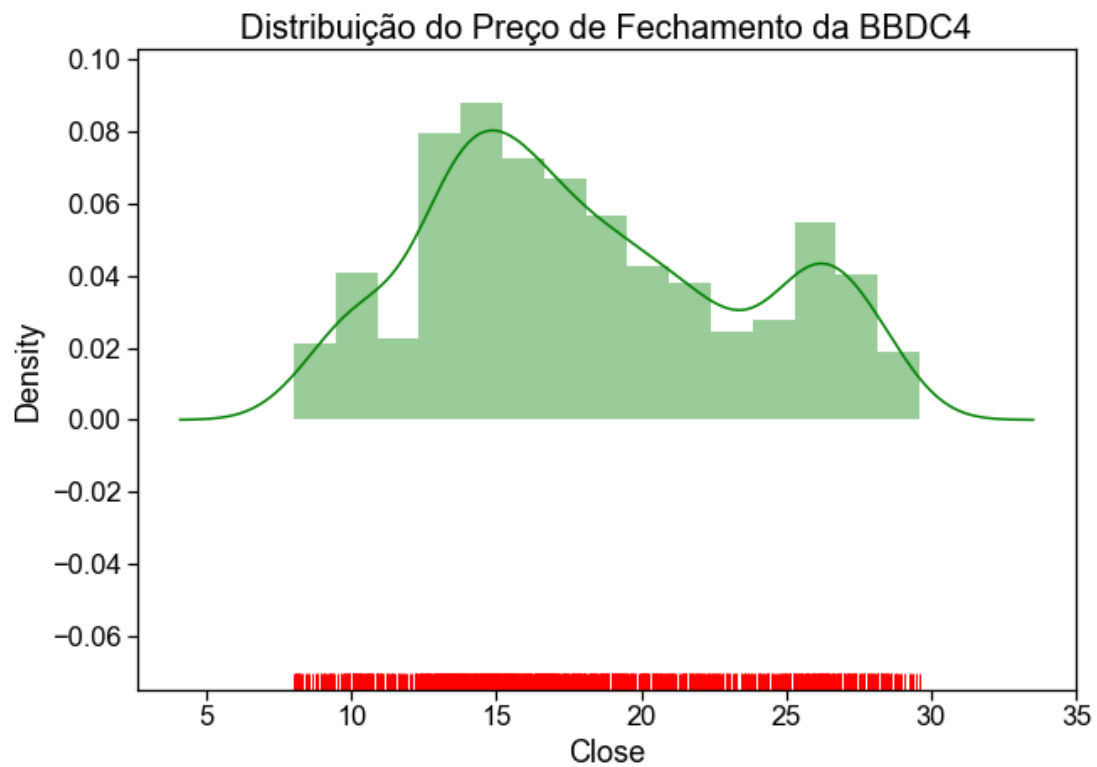
A seguir, verificou-se que não há valores nulos no dataset:

Campo	Valores nulos
High	0
Low	0
Open	0
Close	0
Volume	0
Adj Close	0

Ao plotar o histograma do preço de fechamento da ação, obtém-se o resultado apresentado na figura 7.

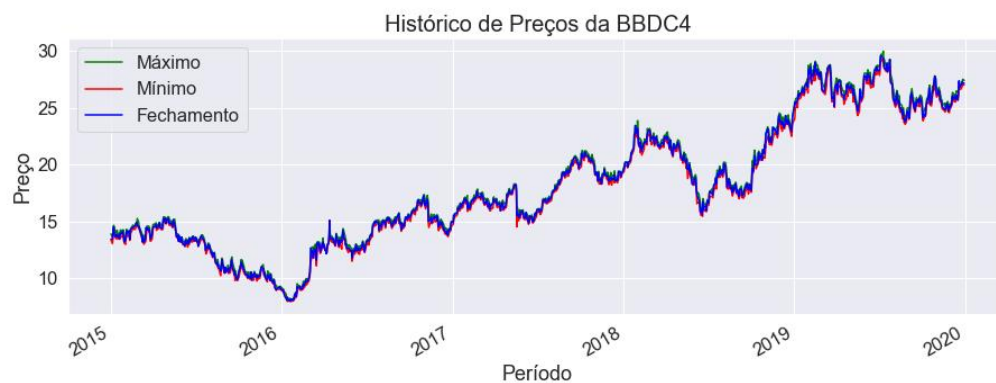
[Digite aqui]

Figura 7: Histograma do preço de fechamento da BBDC4



Também foi feita a plotagem da série histórica de preços contemplando o preço máximo, mínimo e o fechamento, conforme mostra a figura 8.

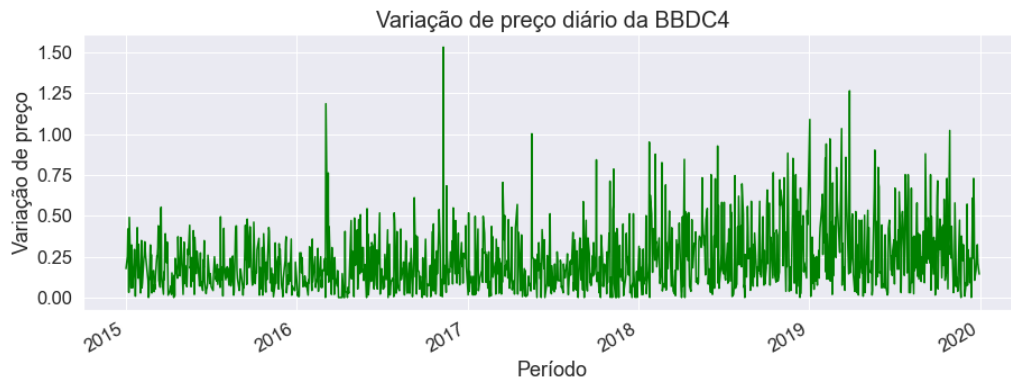
Figura 8: Histórico de preços da BBDC4



[Digite aqui]

Plotando a variação de preço diário (preço de fechamento - preço de abertura), obtém-se o resultado apresentado na figura 9.

Figura 9: Variação de preço diário (preço de fechamento - preço de abertura) da BBDC4

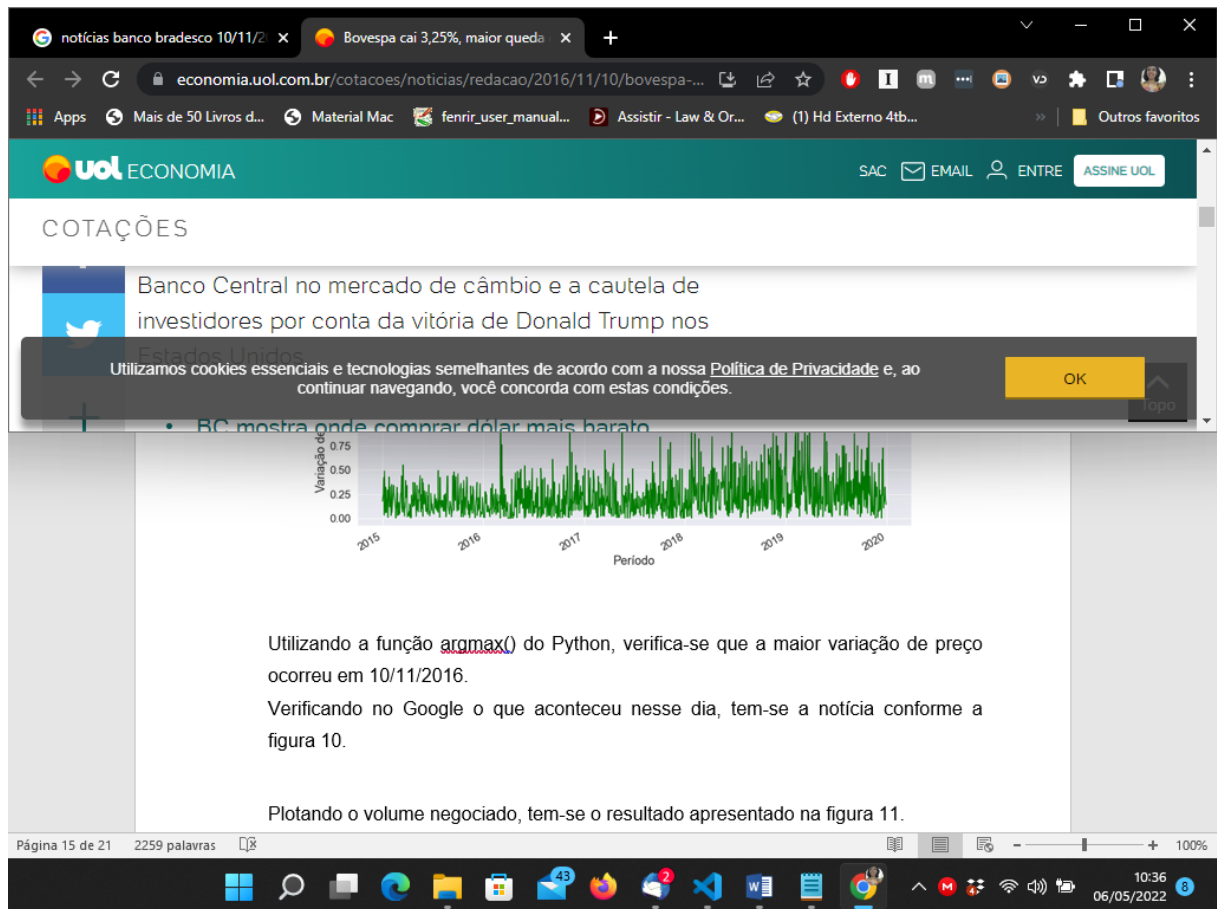


Utilizando a função `argmax()` do Python, verifica-se que a maior variação de preço ocorreu em 10/11/2016.

Verificando no Google o que aconteceu nesse dia, tem-se a notícia conforme a figura 10.

[Digite aqui]

Figura 10: Notícia da BBDC4 encontrada no Google no dia 10/11/2016



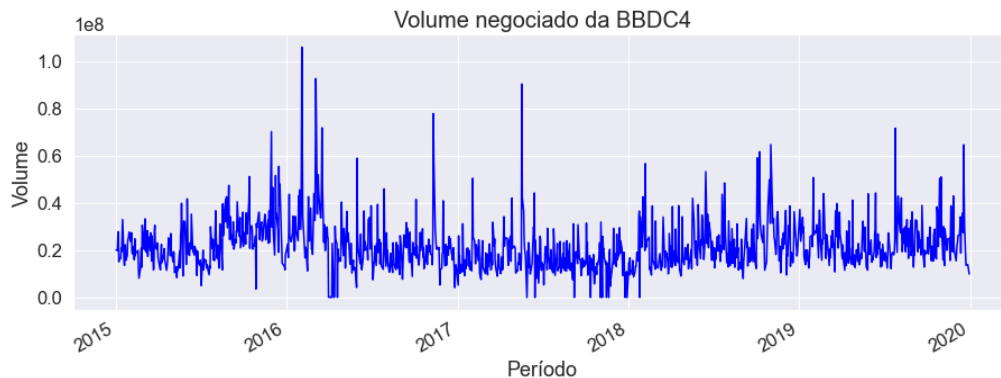
[Notícia da BBDC4 encontrada no Google no dia 10/11/2016](#)

Analisando-se a notícia, percebe-se que o que influenciou essa variação foi a incerteza quanto ao futuro político do então presidente do Brasil Michel Temer, a cautela de investidores por conta da vitória de Donald Trump nos Estados Unidos e a ausência do Banco Central no mercado de câmbio.

Plotando o volume negociado, tem-se o resultado apresentado na figura 11.

[Digite aqui]

Figura 11: Volume negociado da BBDC4



Utilizando a função `argmax()` do Python, verifica-se que o dia com maior volume de negociação foi em 03/02/2016.

Verificando no Google o que aconteceu nesse dia, tem-se a notícia conforme a figura 12.

Figura 12: Notícia da BBDC4 encontrada no Google no dia 03/02/2016

notícias banco bradesco 03/02/2016

Bradesco cancela aumento de capital

G1 - Bradesco cancela aumento de capital

g1.globo.com/economia/negocios/noticia/2016/02/bradesco-cancela-aume...

globo.com g1 ge gshow globoplay

ASSINE JÁ ENTRAR

ECONOMIA

NEGÓCIOS

03/02/2016 10h50 - Atualizado em 03/02/2016 13h20

REUTERS

Bradesco cancela aumento de capital de R\$ 3 bilhões por 'volatilidade'

Ações da instituição financeira subiam mais de 3% nesta quarta-feira (3). No 4º trimestre de 2015, banco teve lucro de R\$ 4,353 bilhões.

Da Reuters

O Conselho de Administração do **Bradesco** decidiu cancelar aumento de capital de R\$ 3 bilhões, citando a volatilidade dos mercados - o que fazia as ações do grupo financeiro subirem mais de 3% nesta quarta-feira (3).

Negócios

veja tudo sobre >

Após ter carro no espaço, Tesla divulga prejuízo recorde no...

10/07/2018

Fox convoca assembleia de acionistas para 10 de julho sobre...

30/05/2018

Bilionário George Soros assume participação em bônus da...

16/05/2018

Toyota abre mais de 800 vagas com criação do

g1.globo.com/economia/negocios/noticia/2016/02/bradesco...

[Digite aqui]

[Notícia da BBDC4 encontrada no Google no dia 03/02/2016](#)

Analisando-se a notícia, percebe-se que o que influenciou esse grande volume de negociação foi o súbito cancelamento de capital por volatilidade dos mercados por parte do banco, fazendo o volume de negociações disparar.

4.3. Análise Exploratória da ITUB4.SA

Após a coleta dos dados, foi realizada uma descrição estatística do dataset, conforme mostra a tabela a seguir:

Cam p o	High	Low	Open	Close	Volume	Adj Close
Co unt	1243.0	1243.0	1243.0	1243.0	1243.0	1243.0
Me an	26.599351 68115151	25.964648 63268483	26.289250 03855821	26.289319 11127853	20962847. 840707965	22.283294 148341554
Std	6.9896115 6720624	6.8754676 61667881	6.9435281 349662095	6.9365539 43996655	10449950. 051949063	7.3689340 170302104
Min	14.072727 20336914	13.824241 638183594	13.939393 043518066	13.981818 199157715	0.0	10.671473 503112793
25 %	19.920109 748840332	19.327548 027038574	19.703028 678894043	19.650136 947631836	14093625. 0 14.984694 004058838	
50 %	26.5	25.833332 061767578	26.200000 762939453	26.173332 21435547	19137750. 0	21.633975 982666016
75 %	33.805000 30517578	33.021665 57312012	33.333332 06176758	33.423334 1217041	25361602. 5	29.709477 424621582
Ma x	39.790000 915527344	38.400001 525878906	38.669998 16894531	39.689998 626708984	104917200 .0	35.422161 10229492

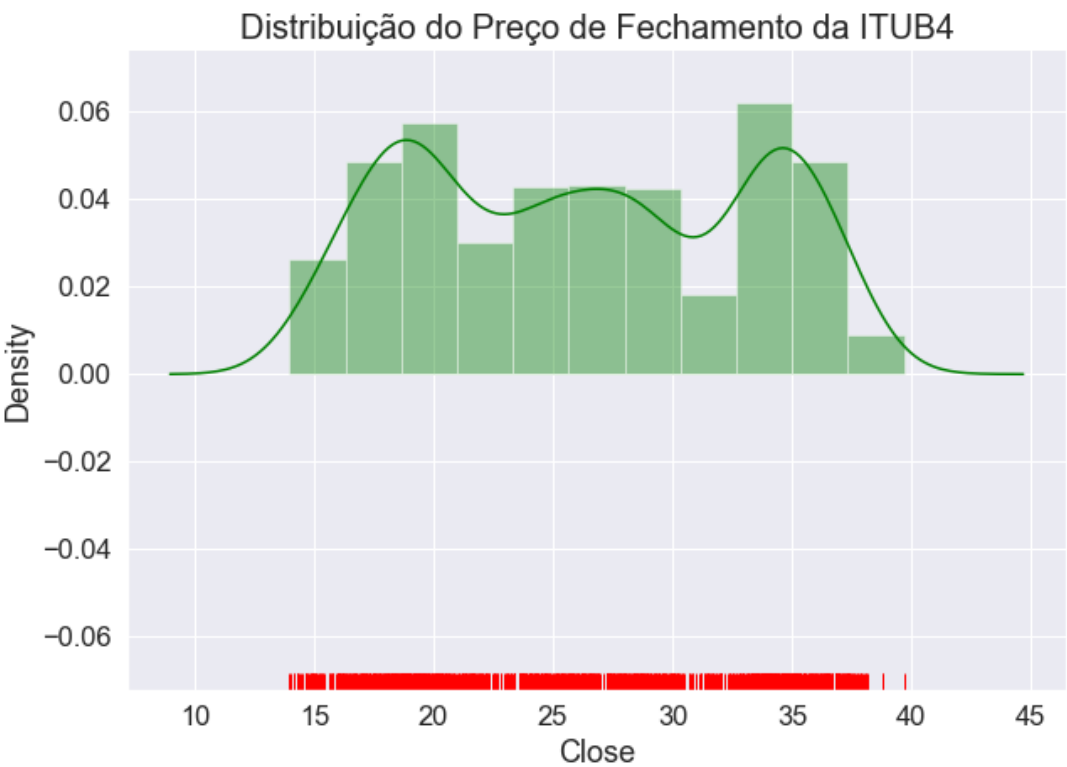
[Digite aqui]

A seguir, verificou-se que não há valores nulos no dataset:

Campo	Valores nulos
High	0
Low	0
Open	0
Close	0
Volume	0
Adj Close	0

Ao plotar o histograma do preço de fechamento da ação, obtém-se o resultado apresentado na figura 13.

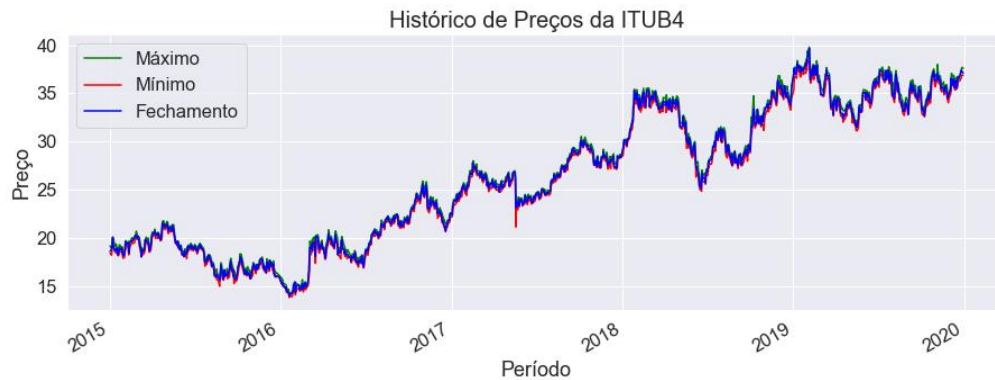
Figura 13: Histograma do preço de fechamento da ITUB4



Também foi feita a plotagem da série histórica de preços contemplando o preço máximo, mínimo e o fechamento, conforme mostra a figura 14.

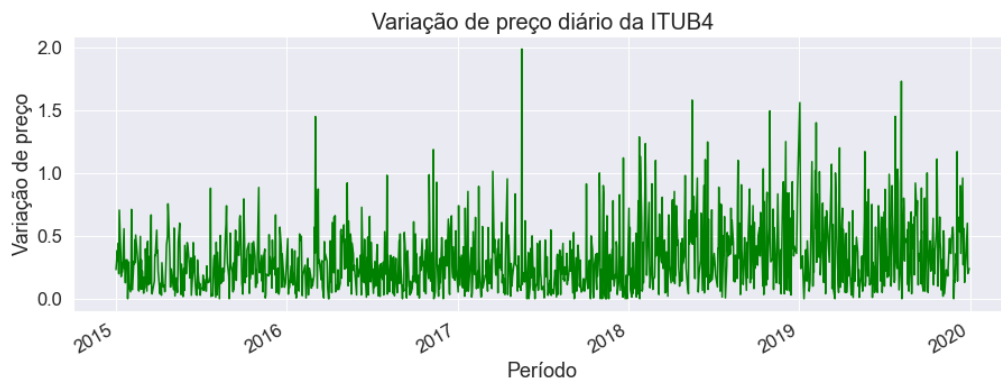
[Digite aqui]

Figura 14: Histórico de preços da ITUB4



Plotando a variação de preço diário (preço de fechamento - preço de abertura), obtém-se o resultado apresentado na figura 15.

Figura 15: Variação de preço diário (preço de fechamento - preço de abertura) da ITUB4

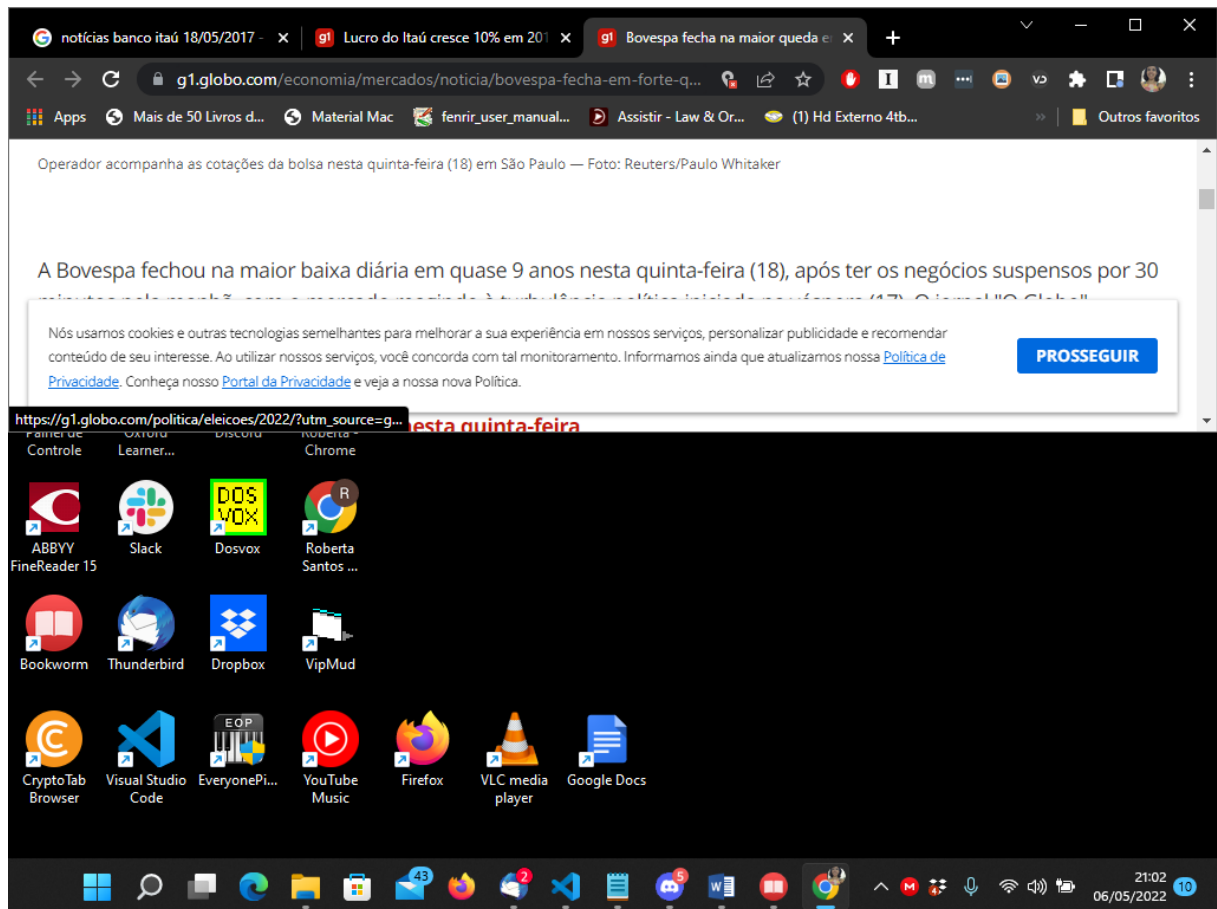


Utilizando a função `argmax()` do Python, verifica-se que a maior variação de preço ocorreu em 18/05/2017.

Verificando no Google o que aconteceu nesse dia, tem-se a notícia conforme a figura 16.

[Digite aqui]

Figura 16: Notícia da ITUB4 encontrada no Google no dia 18/05/2017



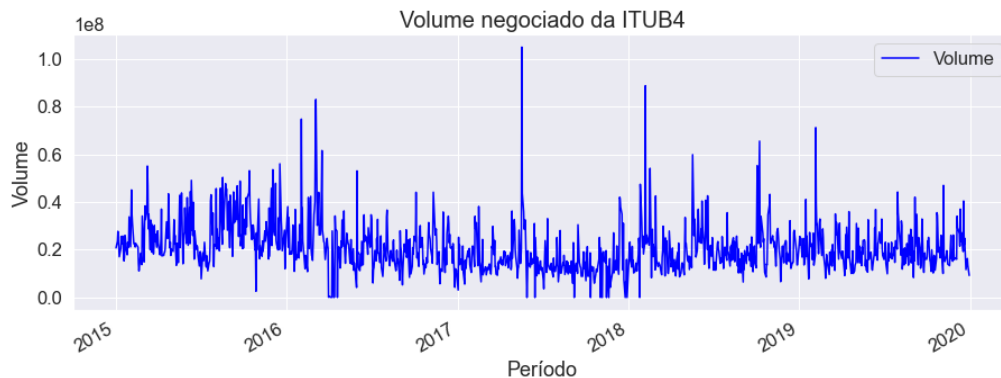
[Notícia da ITUB4 encontrada no Google no dia 18/05/2017](#)

Analisando-se a notícia, percebe-se que o que influenciou essa variação foi a queda generalizada do índice Bovespa, após o presidente da JBS denunciar ter gravado o então presidente do Brasil, Michel Temer, dando aval para comprar silêncio do ex-presidente da Câmara, Eduardo Cunha.

Plotando o volume negociado, tem-se o resultado apresentado na figura 17.

[Digite aqui]

Figura 17: Volume negociado da ITUB4



Utilizando a função `argmax()` do Python, verifica-se que o dia com maior volume de negociação foi em 18/05/2017.

Verificando no Google o que aconteceu nesse dia, tem-se a notícia conforme a figura 16.

Analisando-se a notícia, percebe-se que o que influenciou esse volume de negociação foi a queda generalizada do índice Bovespa devido ao cenário político, conforme explicado anteriormente.

4.4. Análise das Ações em Conjunto

A figura 18 apresenta um gráfico comparativo de desempenho das três ações de forma conjunta. Para isso, foi utilizado o preço ajustado (Adj Close), já que este considera os ajustes de dividendos, splits, inplits, etc. Além disso, o preço ajustado foi normalizado para facilitar a comparação.

[Digite aqui]

Figura 18: Comparativo de retorno das ações BBAS3, BBDC4 e ITUB4

A figura 19 apresenta um gráfico comparativo do log (retorno diário) considerando o valor ajustado das ações.

[Digite aqui]

Figura 19: Comparativo de retorno diário da BBAS3, BBDC4 e ITUB4

Por fim, para cada ação, foram calculados o retorno médio diário, o percentual de retorno anual e o risco, conforme mostra a tabela abaixo:

<i>Ação</i>	<i>Retorno médio diário</i>	<i>Percentual de retorno anual</i>	<i>Risco</i>
<i>BBAS3.SA</i>	<i>0.0009084411865504321</i>	<i>22.7110296637608</i>	<i>0.027639436971218623</i>
<i>BBDC4.SA</i>	<i>0.000775764127565205</i>	<i>19.394103189130124</i>	<i>0.020893431531208543</i>
<i>ITUB4.SA</i>	<i>0.0007601450579285623</i>	<i>19.003626448214057</i>	<i>0.018722353831694908</i>

5. Criação de Modelos de Machine Learning

Esta seção apresenta os modelos preditivos desenvolvidos em linguagem Python para a ação BBAS3, utilizando as bibliotecas Facebook Prophet, Auto-Arima e RNN (Recurrent Neural Network) utilizando a arquitetura LSTM (Long Short Term Memory).

[Digite aqui]

Após a coleta e carregamento dos dados de teste, foi realizada uma descrição estatística do dataset, conforme mostra a tabela a seguir:

Ca mp o	High	Low	Open	Close	Volume	Adj Close
Co unt	248.0	248.0	248.0	248.0	248.0	248.0
Me an	36.101532 313131514	34.746451 65474184	35.466532 314977336	35.372338 610310706	18635531. 85483871	31.155057 568703928
Std	7.2904821 19782365	7.3465991 01258888	7.3477559 17541037	7.3374392 46392954	8003185.6 38362926	6.2239025 8633011
Min	25.0	21.909999 84741211	24.200000 762939453	22.129999 1607666	6435600.0	19.431898 11706543
25 %	31.267499 446868896	30.054999 828338623	30.507500 171661377	30.577500 343322754	13041375. 0	27.119482 99407959
50 %	34.165000 915527344	33.0	33.625	33.444999 69482422	17520050. 0	29.579750 061035156
75 %	37.752498 626708984	35.692500 11444092	36.545000 076293945	36.407499 31335449	22826925. 0	32.349930 76324463
Ma x	53.970001 220703125	53.169998 16894531	53.400001 525878906	53.799999 23706055	73978000. 0	46.584743 49975586

5.1. Modelo Preditivo Com Facebook Prophet

Para esse modelo, o dataset de treinamento foi reduzido em apenas duas colunas, com o objetivo de atender a especificação da biblioteca. Segue abaixo o cabeçalho:

Ds	Y
2015-01-02	22.649999618530273
2015-01-05	22.18000030517578
2015-01-06	22.489999771118164

[Digite aqui]

2015-01-07	23.479999542236328
2015-01-08	23.559999465942383

Onde "ds" representa a data da negociação (date) e "y" o preço de fechamento do dia (close).

Os parâmetros do modelo estão descritos abaixo:

Changepoint_prior_scale: Parâmetro que modula a flexibilidade da seleção automática de ponto de mudança. Valores grandes permitirão muitos pontos de mudança, valores pequenos permitirão poucos pontos de mudança.

Interval_width: Largura dos intervalos de incerteza previstos para a previsão.

Daily_seasonality: Ajusta a sazonalidade diária.

5.2. Modelo Preditivo Com Auto-Arima

Para a utilização da biblioteca AUTO-ARIMA, foi realizado um breve estudo da série temporal, considerando aspectos como sazonalidade, estacionaridade e autocorrelação.

Decompondo-se a série temporal, obtém-se as componentes apresentadas na figura 20.

[Digite aqui]

Figura 20: Decomposição da série temporal da BBAS3

E plotando a autocorrelação, temos o resultado apresentado na figura 21.

[Digite aqui]

Figura 21: Autocorrelação da série temporal da BBAS3

Foi utilizado o modelo AUTO-ARIMA para identificação do melhor modelo de forma automática. A estacionaridade foi verificada utilizando o Teste Dickey-Fuller, obtendo-se o seguinte resultado:

Resultado do Teste Dickey-Fuller:

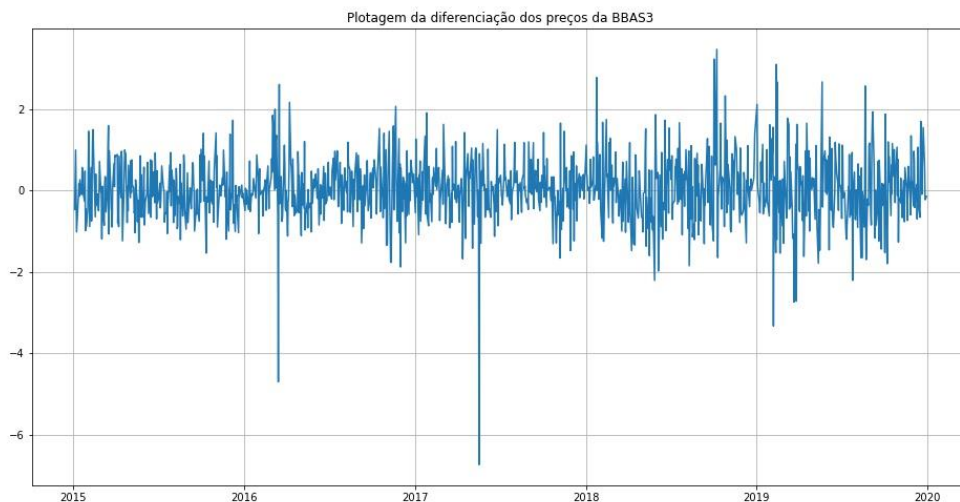
Teste	-0.650803
Valor p	0.859085
Nº de lags	0.000000
Nº de observações	1242.000000
Valor Crítico (1%)	-3.435626
Valor Crítico (5%)	-2.863870
Valor Crítico (10%)	-2.568010

Conforme já era esperado pela análise dos gráficos anteriores, e como o Valor p é muito maior que 0.05, conclui-se que a série não é estacionária. Para torná-la estacionária, poderia ser tentada a aplicação de uma função logarítmica ou a diferenciação da série. Neste caso, foi escolhida a segunda opção.

[Digite aqui]

Assim, aplicando-se uma diferenciação de primeira ordem, obtém-se o resultado da figura 22.

Figura 22: Plotagem da diferenciação de primeira ordem da série temporal



Aparentemente, a série é estacionária. Para confirmar, aplica-se novamente o Teste Dickey-Fuller e obtém-se o seguinte resultado:

Resultado do Teste Dickey-Fuller:

Teste	-34.941323
Valor p	0.000000
Nº de lags	0.000000
Nº de observações	1241.000000
Valor Crítico (1%)	-3.435630
Valor Crítico (5%)	-2.863872
Valor Crítico (10%)	-2.568011

O teste confirma que a série é estacionária, já que o Valor p é muito menor que 0.05.

Plotando os gráficos ACF (Autocorrelation) e PACF (Partial Autocorrelation), tem-se a figura 23.

[Digite aqui]

Figura 23: Gráficos ACF e

PACF

Com estes gráficos, poderiam ser obtidos os componentes p (número de time lags do modelo auto-regressivo) e q (ordem do modelo de média-móvel), caso não fosse utilizada a função AUTO-ARIMA.

Utilizando-se o fato de a série ser estacionária para uma diferenciação de primeira ordem ($d=1$), pode-se utilizar a função AUTO-ARIMA para determinar o melhor modelo que se adequa à série temporal estudada.

O melhor modelo encontrado foi o SARIMAX.

Os principais parâmetros são:

p : Número de time lags do modelo auto-regressivo (AR)

q : Ordem do modelo de média-móvel (MA)

d : Grau de diferenciação

P : Refere-se ao termo auto-regressivo para a parte sazonal

Q : Refere-se ao termo de diferenciação para a parte sazonal

D : Refere-se ao termo da média-móvel para a parte sazonal

5.3. Modelo Preditivo Com RNN (LSTM – Long Short Term Memory)

[Digite aqui]

Para a criação e treinamento do modelo utilizando Redes Neurais Recorrentes – LSTM, o código foi desenvolvido seguindo as seguintes etapas:

1. Carregamento dos dados
2. Normalização dos dados
3. Definição da janela temporal a ser utilizada (60 timesteps)
4. Reshaping do dataset de treinamento
5. Criação do modelo
6. Compilação do modelo
7. Treinamento

O código começa com a inicialização da RNN utilizando a função `Sequential()`. O modelo desenvolvido utiliza duas camadas LSTM com Dropout e uma camada densa de saída com ativação linear. O Dropout serve para reduzir o overfitting, pois ele “desativa” momentaneamente um percentual de neurônios artificiais, neste caso 30%.

A primeira camada ou camada de entrada, possui 128 neurônios do tipo LSTM, a segunda possui 64 e a última possui apenas uma saída linear. A escolha do número de neurônios foi feita de forma empírica e com base na literatura científica.

A compilação utiliza o otimizador Adam, que apresentou bons resultados.

Além disso, foi utilizado como função de perda o erro MSE (Mean Squared Error) e métrica o MAE (Mean Absolute Error).

Para o treinamento, foram utilizadas 100 épocas, este número também foi escolhido de forma empírica. O batch-size é o número de sequências que são treinadas juntas, ou seja, o número de amostras que serão propagadas através da rede neural.

O valor de 32 foi escolhido com base na literatura, que sugere valores como 32, 64, 128, etc.

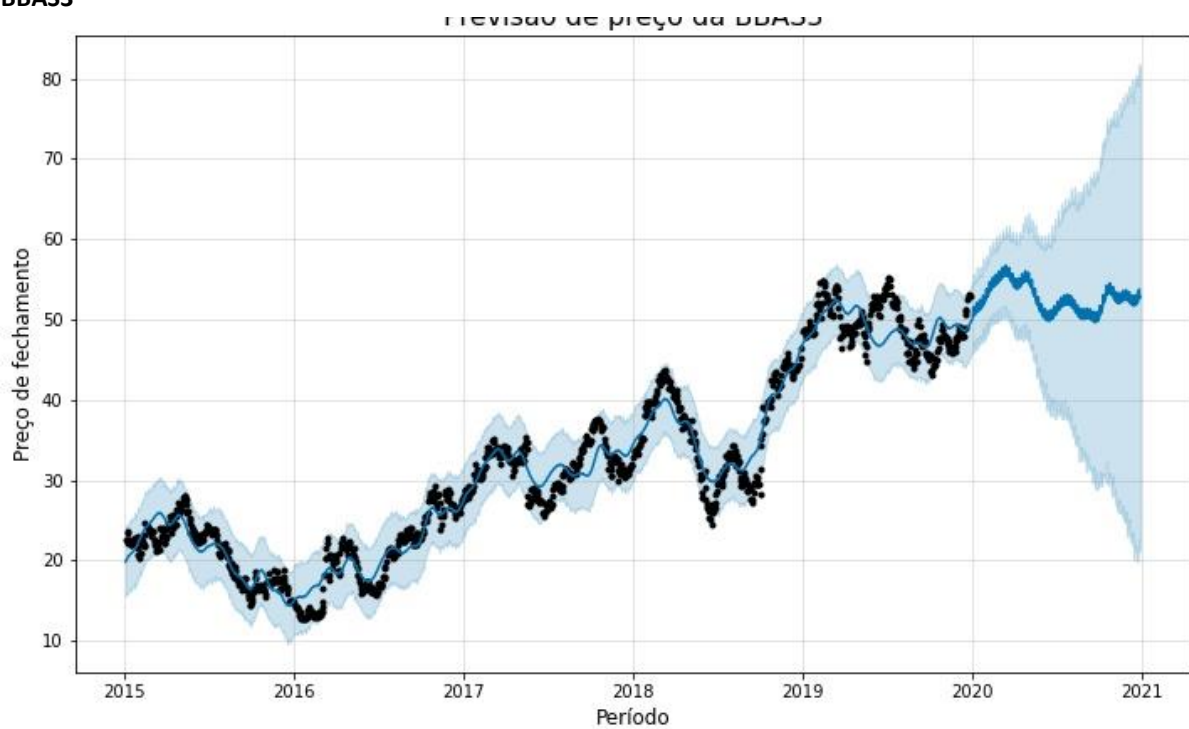
6. Apresentação dos Resultados

6.1. Modelo Preditivo Com Facebook Prophet

A figura 24 mostra o resultado da predição para o ano de 2020. Pode-se perceber que para o ano previsto, o preço de fechamento da BBAS3 apresenta tendência de subida.

[Digite aqui]

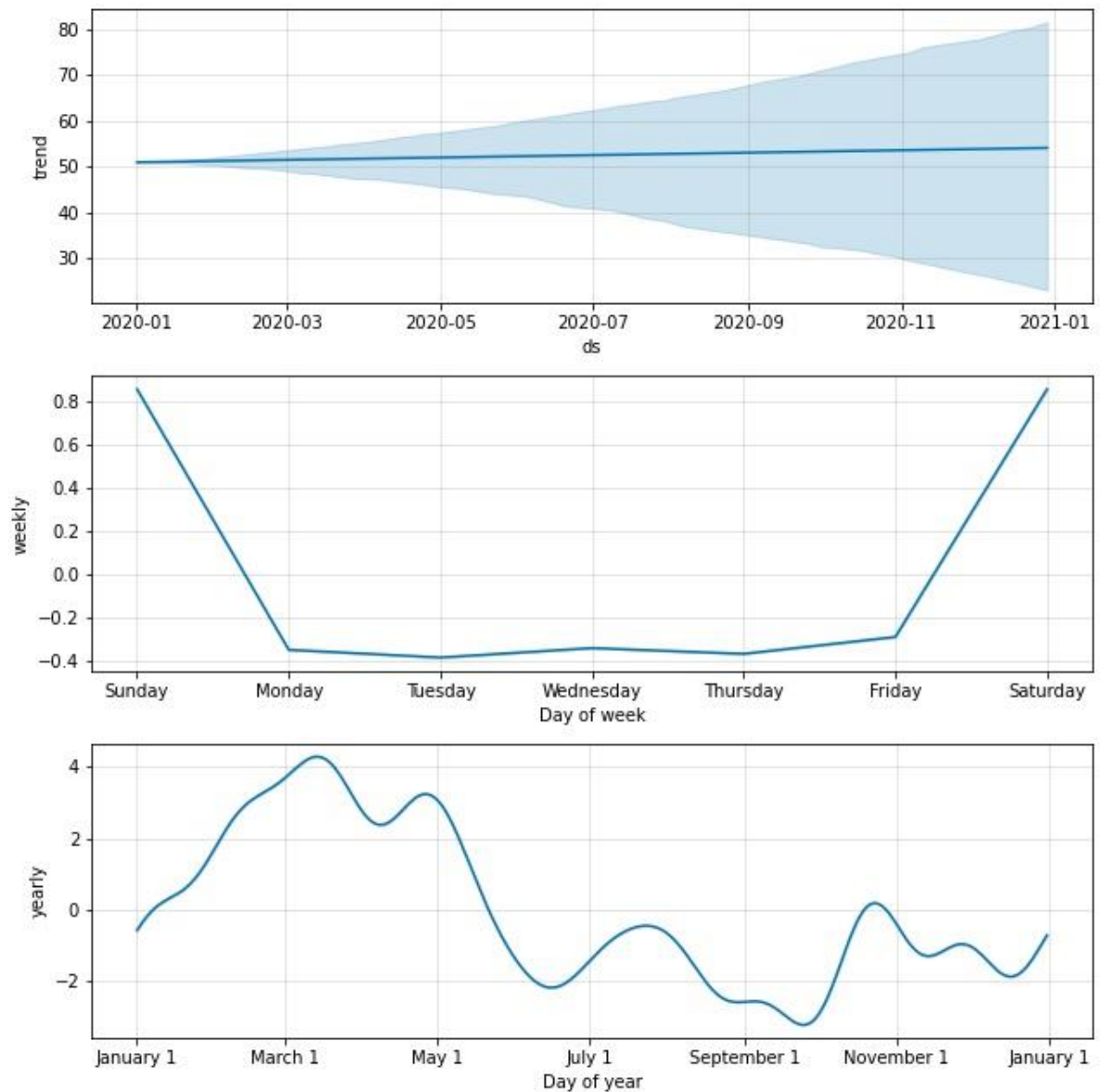
Figura 24: Resultado da previsão de preço da BBAS3



A figura 25 apresenta as componentes decompostas da série temporal prevista. Conforme analisado na figura 24, pode-se confirmar que há uma tendência de subida do preço de fechamento.

[Digite aqui]

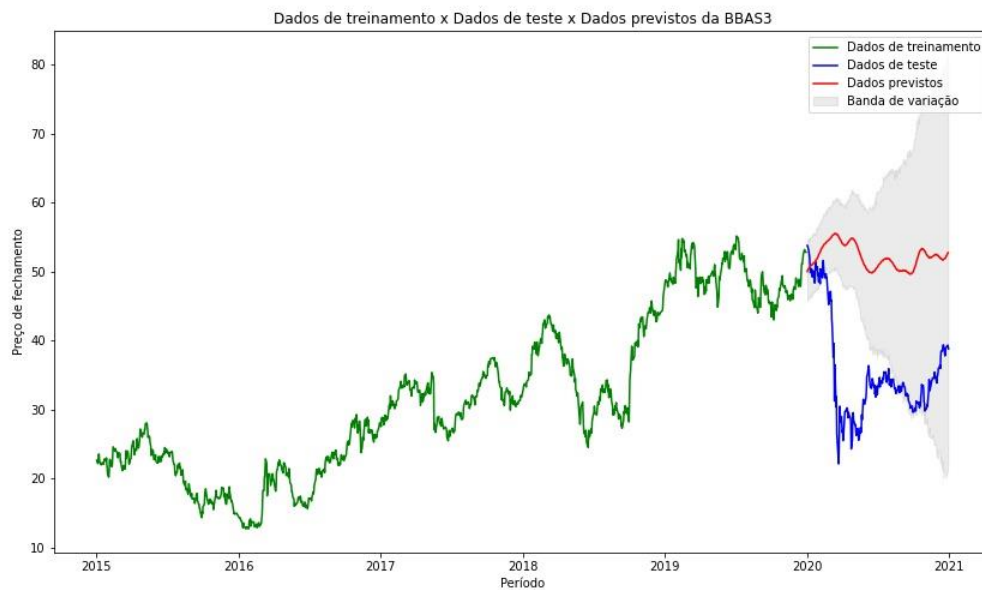
Figura 25: Componentes da previsão de preço da BBAS3



A figura 26 mostra os dados de treinamento, teste, previsão e sua banda de variação feita pelo Prophet para a BBAS3.

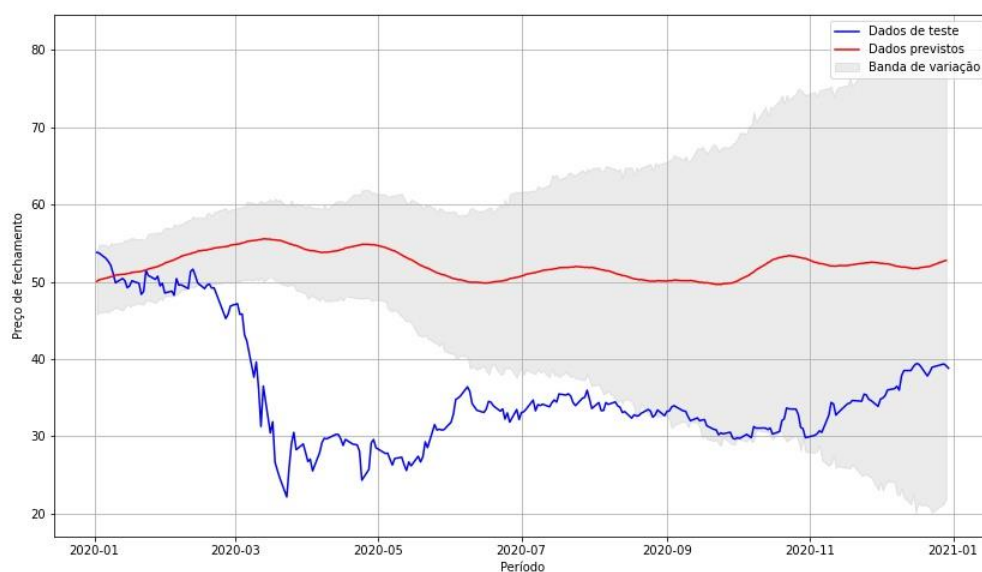
[Digite aqui]

Figura 26: Dados de treinamento, teste e previsão plotados em conjunto



A previsão ficou dentro da banda de variação e acompanha a tendência dos dados de teste. A figura 27 mostra os dados de teste e dados previstos para melhor comparação.

Figura 27: Dados de teste e dados previstos no ano de 2020

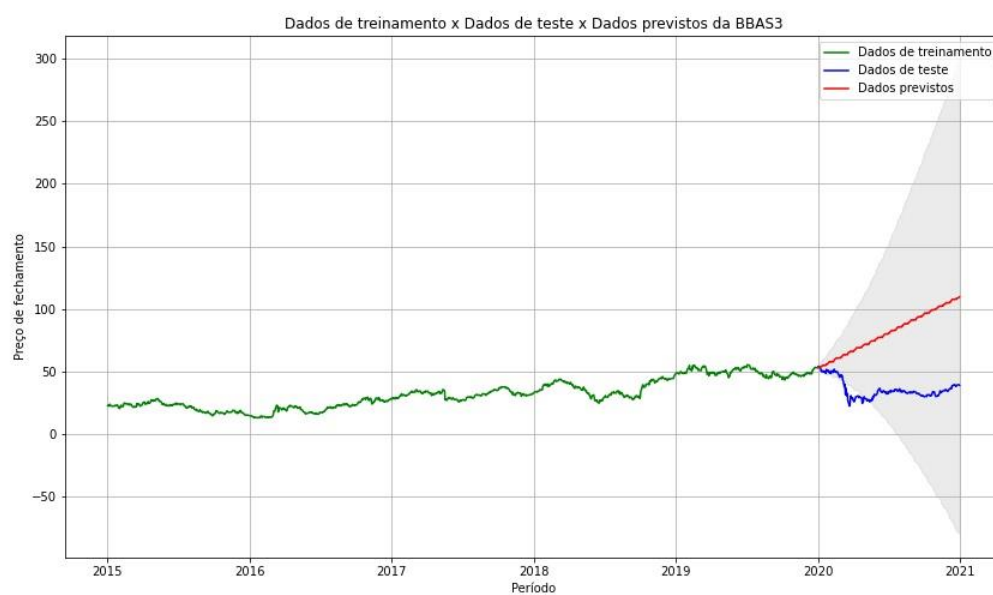


[Digite aqui]

6.2. Modelo Preditivo Com Auto-Arima

A figura 28 mostra o resultado da predição para o ano de 2020. Pode-se perceber que para o ano previsto, o preço de fechamento da BBAS3 apresenta tendência de subida.

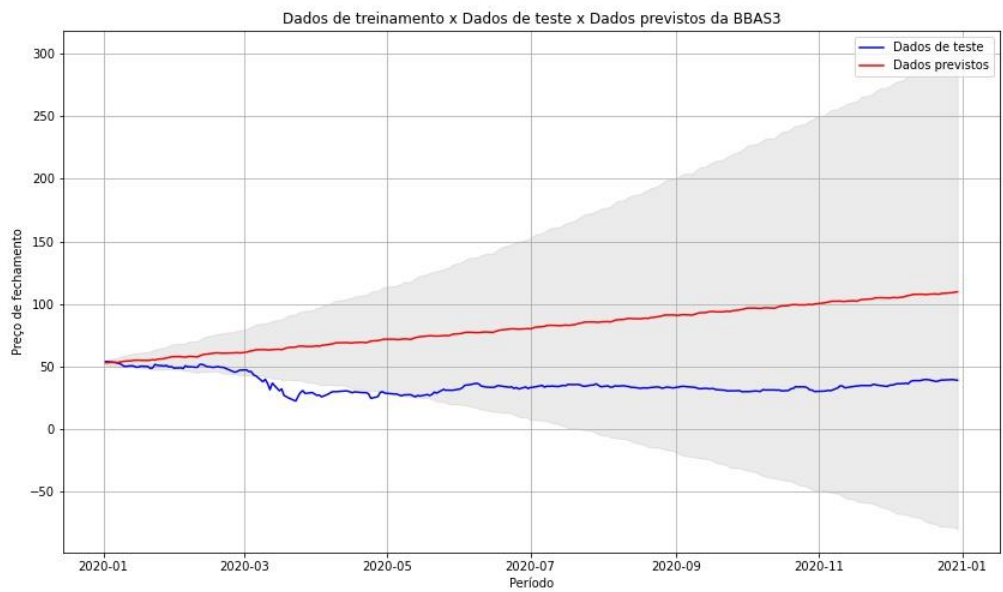
Figura 28: Dados de treinamento, teste e previsão plotados em conjunto



A previsão ficou dentro da banda de variação e acompanha a tendência dos dados de teste. A figura 29 mostra os dados de teste e dados previstos para melhor comparação.

[Digite aqui]

Figura 29: Dados de teste e dados previstos no ano de 2020



Plotagem dos resíduos:

Figura 30: Plotagem dos resí-

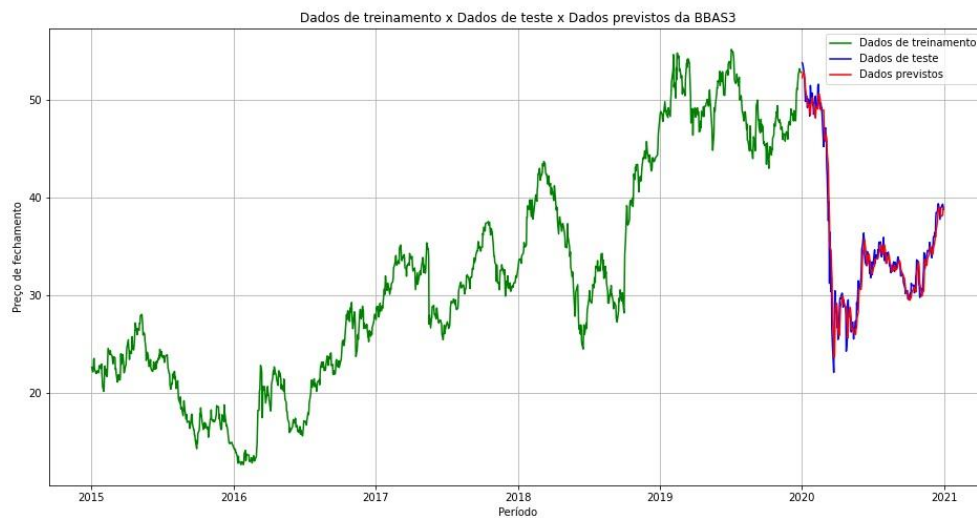
duos

[Digite aqui]

6.3. Modelo Preditivo Com RNN (LSTM – Long Short Term Memory)

A figura 31 mostra o resultado da predição para o ano de 2020. Pode-se perceber que os dados de teste se ajustam de forma muito aproximada com os dados previstos pela RNN.

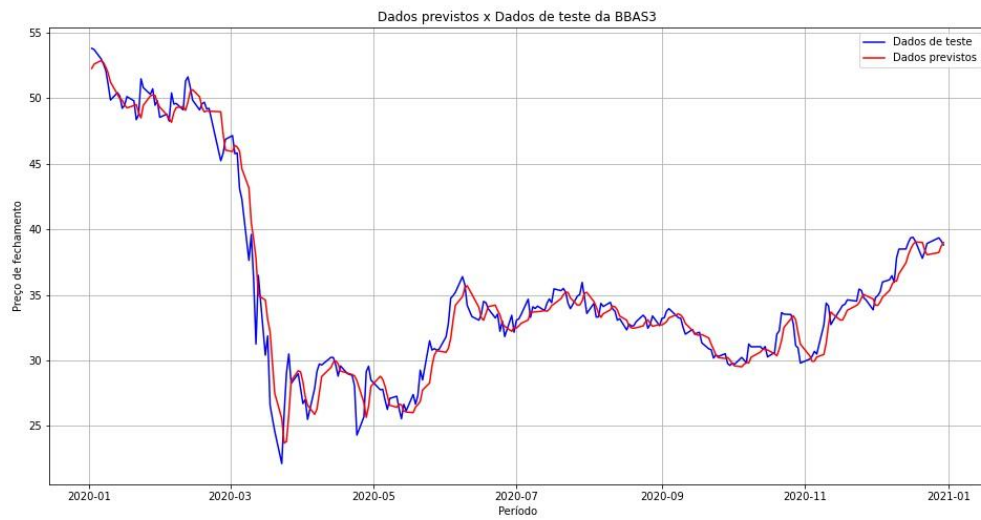
Figura 31: Dados de treinamento, teste e previsão plotados em conjunto



Os dados previstos se ajustam muito bem aos dados de teste. A figura 32 mostra os dados de teste e dados previstos para melhor comparação.

[Digite aqui]

Figura 32: Dados de teste e dados previstos no ano de 2020



7. Links

Link para o vídeo:

Link para o repositório:

[Digite aqui]

REFERÊNCIAS

FACEBOOK Prophet. Disponível em: <https://opensource.facebook.com>. Acesso em: 28 mar. 2022.

KERAS. Disponível em: <https://keras.io/>. Acesso em: 04 mar. 2022.