

bid-ask-spread-predicted

September 13, 2023

```
[27]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import yfinance as yf
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error
from sklearn.linear_model import BayesianRidge
```

```
[28]: datasets = ["BAC"]

for dataset in datasets:
    Ticker = yf.Ticker(dataset)
    data = Ticker.history(start="2019-01-01", end="2023-09-12")
    filename = f"{dataset}_data.csv"
    data.to_csv(filename)
    print(f"Download data for {dataset} and saved as {filename}")
```

Download data for BAC and saved as BAC_data.csv

```
[29]: Ticker = 'BAC'
start_date = '2023-08-10'
end_date = '2023-09-13'
data = yf.download(Ticker, start=start_date, end=end_date)
```

[*****100%*****] 1 of 1 completed

```
[30]: Ticker = 'BAC'
start_date = '2023-08-10'
end_date = '2023-09-13'
```

```
[31]: stock_data = yf.download(Ticker, start=start_date, end=end_date)
```

[*****100%*****] 1 of 1 completed

```
[32]: stock_data.head()
```

```
[32]:
```

	Open	High	Low	Close	Adj Close	Volume
Date						
2023-08-10	31.010000	31.320000	30.799999	30.940001	30.684298	26062600
2023-08-11	30.799999	31.400000	30.750000	31.290001	31.031404	35709900
2023-08-14	31.049999	31.110001	30.750000	30.930000	30.674379	32563700
2023-08-15	30.500000	30.549999	29.840000	29.940001	29.692562	43406100
2023-08-16	29.750000	29.840000	29.209999	29.290001	29.047934	47220400

```
[33]: stock_data['Bid_prices'] = stock_data['Low']
stock_data['Ask_prices'] = stock_data['Open']
```

```
[34]: stock_data['Bid_Ask_Spread'] = stock_data['Ask_prices'] -
stock_data['Bid_prices']
```

```
[35]: stock_data['Bid_Ask_Spread'].mean()
```

```
[35]: 0.2739130932351817
```

```
[36]: stock_data
```

```
[36]:
```

	Open	High	Low	Close	Adj Close	Volume	\
Date							
2023-08-10	31.010000	31.320000	30.799999	30.940001	30.684298	26062600	
2023-08-11	30.799999	31.400000	30.750000	31.290001	31.031404	35709900	
2023-08-14	31.049999	31.110001	30.750000	30.930000	30.674379	32563700	
2023-08-15	30.500000	30.549999	29.840000	29.940001	29.692562	43406100	
2023-08-16	29.750000	29.840000	29.209999	29.290001	29.047934	47220400	
2023-08-17	29.440001	29.540001	29.110001	29.280001	29.038017	37202200	
2023-08-18	29.040001	29.350000	28.980000	29.110001	28.869421	31418000	
2023-08-21	29.219999	29.299999	28.799999	29.150000	28.909090	34959000	
2023-08-22	29.129999	29.240000	28.420000	28.440001	28.204958	37586400	
2023-08-23	28.450001	28.510000	28.180000	28.450001	28.214876	44686600	
2023-08-24	28.400000	28.790001	28.370001	28.620001	28.383471	31934800	
2023-08-25	28.639999	28.790001	28.299999	28.500000	28.264462	34253200	
2023-08-28	28.690001	29.000000	28.570000	28.760000	28.522314	33075200	
2023-08-29	28.889999	29.260000	28.719999	29.170000	28.928925	30428200	
2023-08-30	29.219999	29.270000	28.930000	29.040001	28.800001	33366400	
2023-08-31	28.930000	28.969999	28.530001	28.670000	28.670000	37243000	
2023-09-01	28.760000	29.129999	28.760000	28.980000	28.980000	35192000	
2023-09-05	28.940001	29.120001	28.500000	28.650000	28.650000	46543600	
2023-09-06	28.400000	28.559999	28.120001	28.389999	28.389999	45022600	
2023-09-07	28.410000	28.549999	28.059999	28.129999	28.129999	35023600	
2023-09-08	28.120001	28.379999	27.959999	28.360001	28.360001	36513900	
2023-09-11	28.510000	28.760000	28.400000	28.480000	28.480000	34168300	
2023-09-12	28.459999	29.139999	28.400000	28.969999	28.969999	42273200	

Bid_prices Ask_prices Bid_Ask_Spread

Date			
2023-08-10	30.799999	31.010000	0.210001
2023-08-11	30.750000	30.799999	0.049999
2023-08-14	30.750000	31.049999	0.299999
2023-08-15	29.840000	30.500000	0.660000
2023-08-16	29.209999	29.750000	0.540001
2023-08-17	29.110001	29.440001	0.330000
2023-08-18	28.980000	29.040001	0.060001
2023-08-21	28.799999	29.219999	0.420000
2023-08-22	28.420000	29.129999	0.709999
2023-08-23	28.180000	28.450001	0.270000
2023-08-24	28.370001	28.400000	0.029999
2023-08-25	28.299999	28.639999	0.340000
2023-08-28	28.570000	28.690001	0.120001
2023-08-29	28.719999	28.889999	0.170000
2023-08-30	28.930000	29.219999	0.289999
2023-08-31	28.530001	28.930000	0.400000
2023-09-01	28.760000	28.760000	0.000000
2023-09-05	28.500000	28.940001	0.440001
2023-09-06	28.120001	28.400000	0.279999
2023-09-07	28.059999	28.410000	0.350000
2023-09-08	27.959999	28.120001	0.160002
2023-09-11	28.400000	28.510000	0.110001
2023-09-12	28.400000	28.459999	0.059999

```
[66]: stock_data['Previous_Close'] = stock_data['Close'].pct_change()
stock_data['Previous_Close'] = stock_data['Close'].dropna()
```

```
[67]: stock_data.head()
```

```
[67]:
```

	Open	High	Low	Close	Adj Close	Volume \
Date						
2023-08-10	31.010000	31.320000	30.799999	30.940001	30.684298	26062600
2023-08-11	30.799999	31.400000	30.750000	31.290001	31.031404	35709900
2023-08-14	31.049999	31.110001	30.750000	30.930000	30.674379	32563700
2023-08-15	30.500000	30.549999	29.840000	29.940001	29.692562	43406100
2023-08-16	29.750000	29.840000	29.209999	29.290001	29.047934	47220400

	Bid_prices	Ask_prices	Bid_Ask_Spread	Previous_Close
Date				
2023-08-10	30.799999	31.010000	0.210001	30.940001
2023-08-11	30.750000	30.799999	0.049999	31.290001
2023-08-14	30.750000	31.049999	0.299999	30.930000
2023-08-15	29.840000	30.500000	0.660000	29.940001
2023-08-16	29.209999	29.750000	0.540001	29.290001

```
[68]: stock_data.dropna()
```

[68]:

	Open	High	Low	Close	Adj Close	Volume \
Date						
2023-08-10	31.010000	31.320000	30.799999	30.940001	30.684298	26062600
2023-08-11	30.799999	31.400000	30.750000	31.290001	31.031404	35709900
2023-08-14	31.049999	31.110001	30.750000	30.930000	30.674379	32563700
2023-08-15	30.500000	30.549999	29.840000	29.940001	29.692562	43406100
2023-08-16	29.750000	29.840000	29.209999	29.290001	29.047934	47220400
2023-08-17	29.440001	29.540001	29.110001	29.280001	29.038017	37202200
2023-08-18	29.040001	29.350000	28.980000	29.110001	28.869421	31418000
2023-08-21	29.219999	29.299999	28.799999	29.150000	28.909090	34959000
2023-08-22	29.129999	29.240000	28.420000	28.440001	28.204958	37586400
2023-08-23	28.450001	28.510000	28.180000	28.450001	28.214876	44686600
2023-08-24	28.400000	28.790001	28.370001	28.620001	28.383471	31934800
2023-08-25	28.639999	28.790001	28.299999	28.500000	28.264462	34253200
2023-08-28	28.690001	29.000000	28.570000	28.760000	28.522314	33075200
2023-08-29	28.889999	29.260000	28.719999	29.170000	28.928925	30428200
2023-08-30	29.219999	29.270000	28.930000	29.040001	28.800001	33366400
2023-08-31	28.930000	28.969999	28.530001	28.670000	28.670000	37243000
2023-09-01	28.760000	29.129999	28.760000	28.980000	28.980000	35192000
2023-09-05	28.940001	29.120001	28.500000	28.650000	28.650000	46543600
2023-09-06	28.400000	28.559999	28.120001	28.389999	28.389999	45022600
2023-09-07	28.410000	28.549999	28.059999	28.129999	28.129999	35023600
2023-09-08	28.120001	28.379999	27.959999	28.360001	28.360001	36513900
2023-09-11	28.510000	28.760000	28.400000	28.480000	28.480000	34168300
2023-09-12	28.459999	29.139999	28.400000	28.969999	28.969999	42273200

	Bid_prices	Ask_prices	Bid_Ask_Spread	Previous_Close
Date				
2023-08-10	30.799999	31.010000	0.210001	30.940001
2023-08-11	30.750000	30.799999	0.049999	31.290001
2023-08-14	30.750000	31.049999	0.299999	30.930000
2023-08-15	29.840000	30.500000	0.660000	29.940001
2023-08-16	29.209999	29.750000	0.540001	29.290001
2023-08-17	29.110001	29.440001	0.330000	29.280001
2023-08-18	28.980000	29.040001	0.060001	29.110001
2023-08-21	28.799999	29.219999	0.420000	29.150000
2023-08-22	28.420000	29.129999	0.709999	28.440001
2023-08-23	28.180000	28.450001	0.270000	28.450001
2023-08-24	28.370001	28.400000	0.029999	28.620001
2023-08-25	28.299999	28.639999	0.340000	28.500000
2023-08-28	28.570000	28.690001	0.120001	28.760000
2023-08-29	28.719999	28.889999	0.170000	29.170000
2023-08-30	28.930000	29.219999	0.289999	29.040001
2023-08-31	28.530001	28.930000	0.400000	28.670000
2023-09-01	28.760000	28.760000	0.000000	28.980000
2023-09-05	28.500000	28.940001	0.440001	28.650000
2023-09-06	28.120001	28.400000	0.279999	28.389999

2023-09-07	28.059999	28.410000	0.350000	28.129999
2023-09-08	27.959999	28.120001	0.160002	28.360001
2023-09-11	28.400000	28.510000	0.110001	28.480000
2023-09-12	28.400000	28.459999	0.059999	28.969999

```
[69]: stock_data.dropna()
```

```
[69]:
```

	Open	High	Low	Close	Adj Close	Volume \
Date						
2023-08-10	31.010000	31.320000	30.799999	30.940001	30.684298	26062600
2023-08-11	30.799999	31.400000	30.750000	31.290001	31.031404	35709900
2023-08-14	31.049999	31.110001	30.750000	30.930000	30.674379	32563700
2023-08-15	30.500000	30.549999	29.840000	29.940001	29.692562	43406100
2023-08-16	29.750000	29.840000	29.209999	29.290001	29.047934	47220400
2023-08-17	29.440001	29.540001	29.110001	29.280001	29.038017	37202200
2023-08-18	29.040001	29.350000	28.980000	29.110001	28.869421	31418000
2023-08-21	29.219999	29.299999	28.799999	29.150000	28.909090	34959000
2023-08-22	29.129999	29.240000	28.420000	28.440001	28.204958	37586400
2023-08-23	28.450001	28.510000	28.180000	28.450001	28.214876	44686600
2023-08-24	28.400000	28.790001	28.370001	28.620001	28.383471	31934800
2023-08-25	28.639999	28.790001	28.299999	28.500000	28.264462	34253200
2023-08-28	28.690001	29.000000	28.570000	28.760000	28.522314	33075200
2023-08-29	28.889999	29.260000	28.719999	29.170000	28.928925	30428200
2023-08-30	29.219999	29.270000	28.930000	29.040001	28.800001	33366400
2023-08-31	28.930000	28.969999	28.530001	28.670000	28.670000	37243000
2023-09-01	28.760000	29.129999	28.760000	28.980000	28.980000	35192000
2023-09-05	28.940001	29.120001	28.500000	28.650000	28.650000	46543600
2023-09-06	28.400000	28.559999	28.120001	28.389999	28.389999	45022600
2023-09-07	28.410000	28.549999	28.059999	28.129999	28.129999	35023600
2023-09-08	28.120001	28.379999	27.959999	28.360001	28.360001	36513900
2023-09-11	28.510000	28.760000	28.400000	28.480000	28.480000	34168300
2023-09-12	28.459999	29.139999	28.400000	28.969999	28.969999	42273200

	Bid_prices	Ask_prices	Bid_Ask_Spread	Previous_Close
Date				
2023-08-10	30.799999	31.010000	0.210001	30.940001
2023-08-11	30.750000	30.799999	0.049999	31.290001
2023-08-14	30.750000	31.049999	0.299999	30.930000
2023-08-15	29.840000	30.500000	0.660000	29.940001
2023-08-16	29.209999	29.750000	0.540001	29.290001
2023-08-17	29.110001	29.440001	0.330000	29.280001
2023-08-18	28.980000	29.040001	0.060001	29.110001
2023-08-21	28.799999	29.219999	0.420000	29.150000
2023-08-22	28.420000	29.129999	0.709999	28.440001
2023-08-23	28.180000	28.450001	0.270000	28.450001
2023-08-24	28.370001	28.400000	0.029999	28.620001
2023-08-25	28.299999	28.639999	0.340000	28.500000

2023-08-28	28.570000	28.690001	0.120001	28.760000
2023-08-29	28.719999	28.889999	0.170000	29.170000
2023-08-30	28.930000	29.219999	0.289999	29.040001
2023-08-31	28.530001	28.930000	0.400000	28.670000
2023-09-01	28.760000	28.760000	0.000000	28.980000
2023-09-05	28.500000	28.940001	0.440001	28.650000
2023-09-06	28.120001	28.400000	0.279999	28.389999
2023-09-07	28.059999	28.410000	0.350000	28.129999
2023-09-08	27.959999	28.120001	0.160002	28.360001
2023-09-11	28.400000	28.510000	0.110001	28.480000
2023-09-12	28.400000	28.459999	0.059999	28.969999

```
[70]: stock_data.describe()
```

```
[70]:
```

	Open	High	Low	Close	Adj Close	Volume \
count	23.000000	23.000000	23.000000	23.000000	23.000000	2.300000e+01
mean	29.163478	29.386087	28.889565	29.110435	28.952005	3.677621e+07
std	0.878271	0.884065	0.853833	0.870556	0.817627	5.575280e+06
min	28.120001	28.379999	27.959999	28.129999	28.129999	2.606260e+07
25%	28.485000	28.790001	28.385000	28.490000	28.386735	3.322080e+07
50%	28.930000	29.139999	28.570000	28.969999	28.800001	3.519200e+07
75%	29.330000	29.445001	29.045000	29.225000	29.009008	3.992980e+07
max	31.049999	31.400000	30.799999	31.290001	31.031404	4.722040e+07

	Bid_prices	Ask_prices	Bid_Ask_Spread	Previous_Close
count	23.000000	23.000000	23.000000	23.000000
mean	28.889565	29.163478	0.273913	29.110435
std	0.853833	0.878271	0.195837	0.870556
min	27.959999	28.120001	0.000000	28.129999
25%	28.385000	28.485000	0.115001	28.490000
50%	28.570000	28.930000	0.279999	28.969999
75%	29.045000	29.330000	0.375000	29.225000
max	30.799999	31.049999	0.709999	31.290001

```
[71]: stock_data.corr()
```

```
[71]:
```

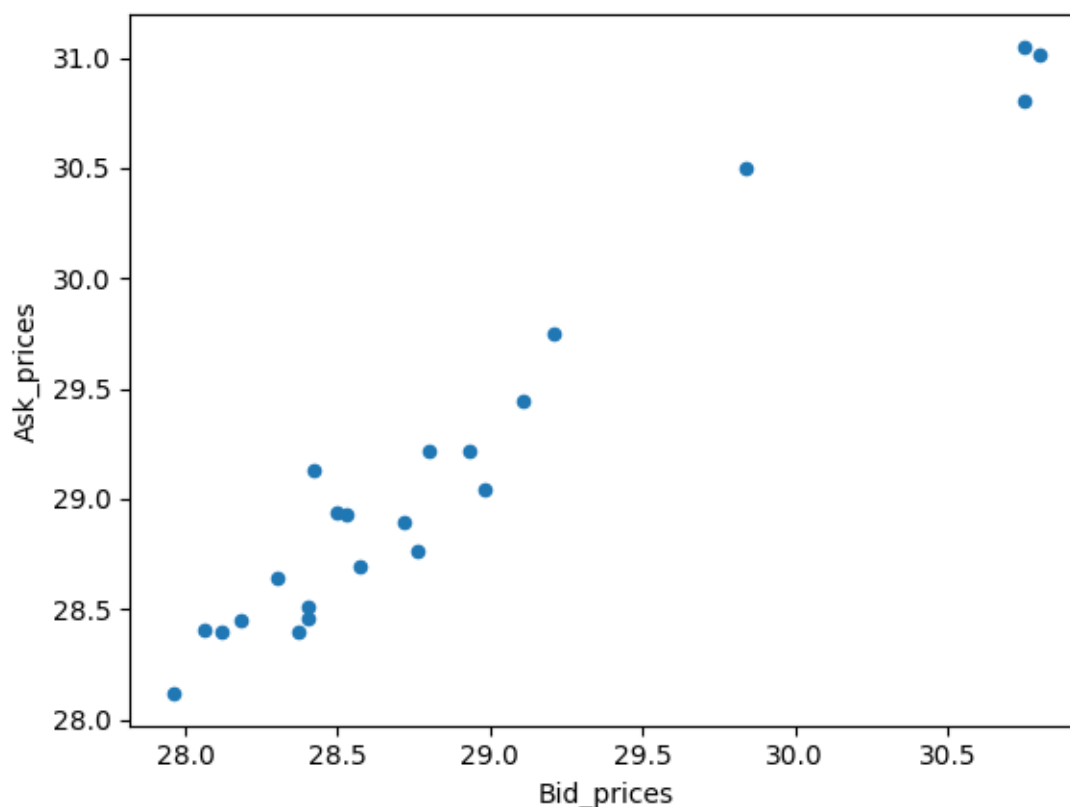
	Open	High	Low	Close	Adj Close	Volume \
Open	1.000000	0.980192	0.974827	0.940224	0.921543	-0.182750
High	0.980192	1.000000	0.988046	0.976932	0.965348	-0.225755
Low	0.974827	0.988046	1.000000	0.985133	0.973031	-0.291881
Close	0.940224	0.976932	0.985133	1.000000	0.992109	-0.286162
Adj Close	0.921543	0.965348	0.973031	0.992109	1.000000	-0.260168
Volume	-0.182750	-0.225755	-0.291881	-0.286162	-0.260168	1.000000
Bid_prices	0.974827	0.988046	1.000000	0.985133	0.973031	-0.291881
Ask_prices	1.000000	0.980192	0.974827	0.940224	0.921543	-0.182750
Bid_Ask_Spread	0.234539	0.088070	0.011890	-0.078476	-0.109487	0.452997
Previous_Close	0.940224	0.976932	0.985133	1.000000	0.992109	-0.286162

	Bid_prices	Ask_prices	Bid_Ask_Spread	Previous_Close
Open	0.974827	1.000000	0.234539	0.940224
High	0.988046	0.980192	0.088070	0.976932
Low	1.000000	0.974827	0.011890	0.985133
Close	0.985133	0.940224	-0.078476	1.000000
Adj Close	0.973031	0.921543	-0.109487	0.992109
Volume	-0.291881	-0.182750	0.452997	-0.286162
Bid_prices	1.000000	0.974827	0.011890	0.985133
Ask_prices	0.974827	1.000000	0.234539	0.940224
Bid_Ask_Spread	0.011890	0.234539	1.000000	-0.078476
Previous_Close	0.985133	0.940224	-0.078476	1.000000

```
[72]: X = stock_data[['Previous_Close']]
      y = stock_data['Bid_Ask_Spread']
```

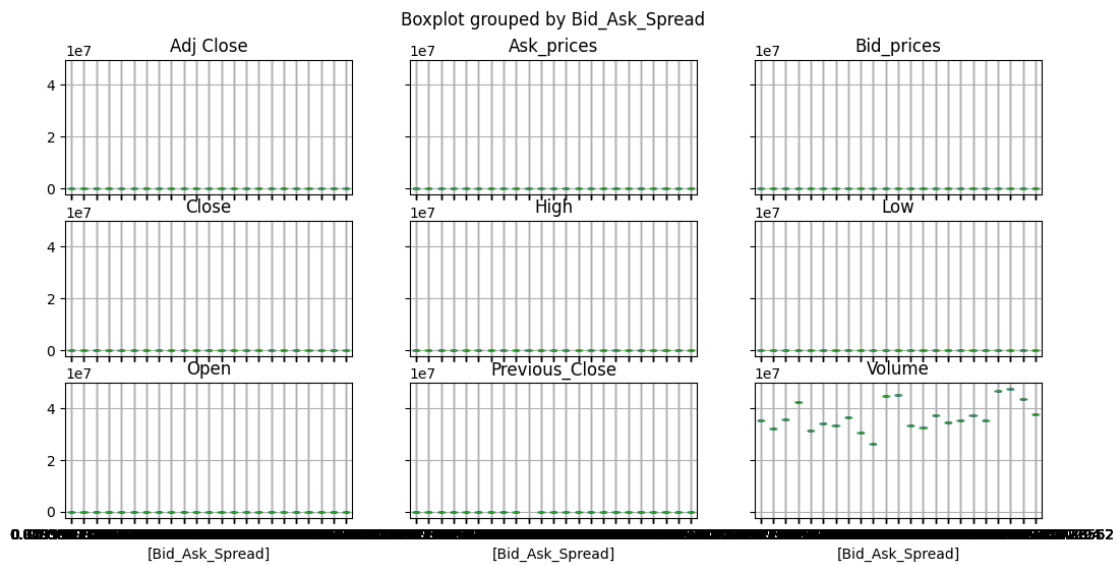
```
[61]: stock_data.plot(kind="scatter", x='Bid_prices', y = 'Ask_prices')
```

```
[61]: <Axes: xlabel='Bid_prices', ylabel='Ask_prices'>
```



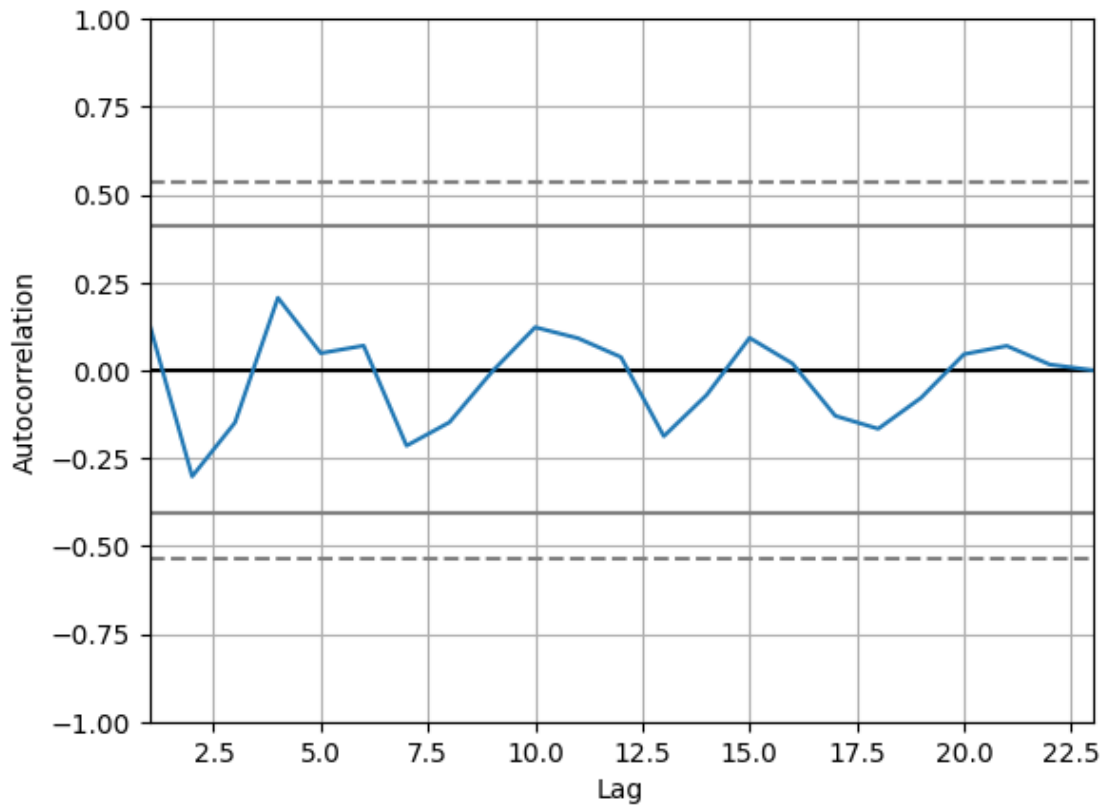
```
[62]: stock_data.boxplot(by='Bid_Ask_Spread', figsize=(12,6))
```

```
[62]: array([[<Axes: title={'center': 'Adj Close'}, xlabel=' [Bid_Ask_Spread] '>,
  <Axes: title={'center': 'Ask_prices'}, xlabel=' [Bid_Ask_Spread] '>,
  <Axes: title={'center': 'Bid_prices'}, xlabel=' [Bid_Ask_Spread] '>],
  [<Axes: title={'center': 'Close'}, xlabel=' [Bid_Ask_Spread] '>,
  <Axes: title={'center': 'High'}, xlabel=' [Bid_Ask_Spread] '>,
  <Axes: title={'center': 'Low'}, xlabel=' [Bid_Ask_Spread] '>],
  [<Axes: title={'center': 'Open'}, xlabel=' [Bid_Ask_Spread] '>,
  <Axes: title={'center': 'Previous_Close'}, xlabel=' [Bid_Ask_Spread] '>,
  <Axes: title={'center': 'Volume'}, xlabel=' [Bid_Ask_Spread] '>]],
  dtype=object)
```



```
[64]: from pandas.plotting import autocorrelation_plot
autocorrelation_plot(stock_data['Bid_Ask_Spread'])
```

```
[64]: <Axes: xlabel='Lag', ylabel='Autocorrelation'>
```

```
[73]: X_train , X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
↳ random_state=42)
```

```
[74]: model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)
```

```
[74]: RandomForestRegressor(random_state=42)
```

```
[75]: y_pred = model.predict(X_test)
```

```
[76]: mse = mean_squared_error(y_test, y_pred)
```

```
[77]: mse
```

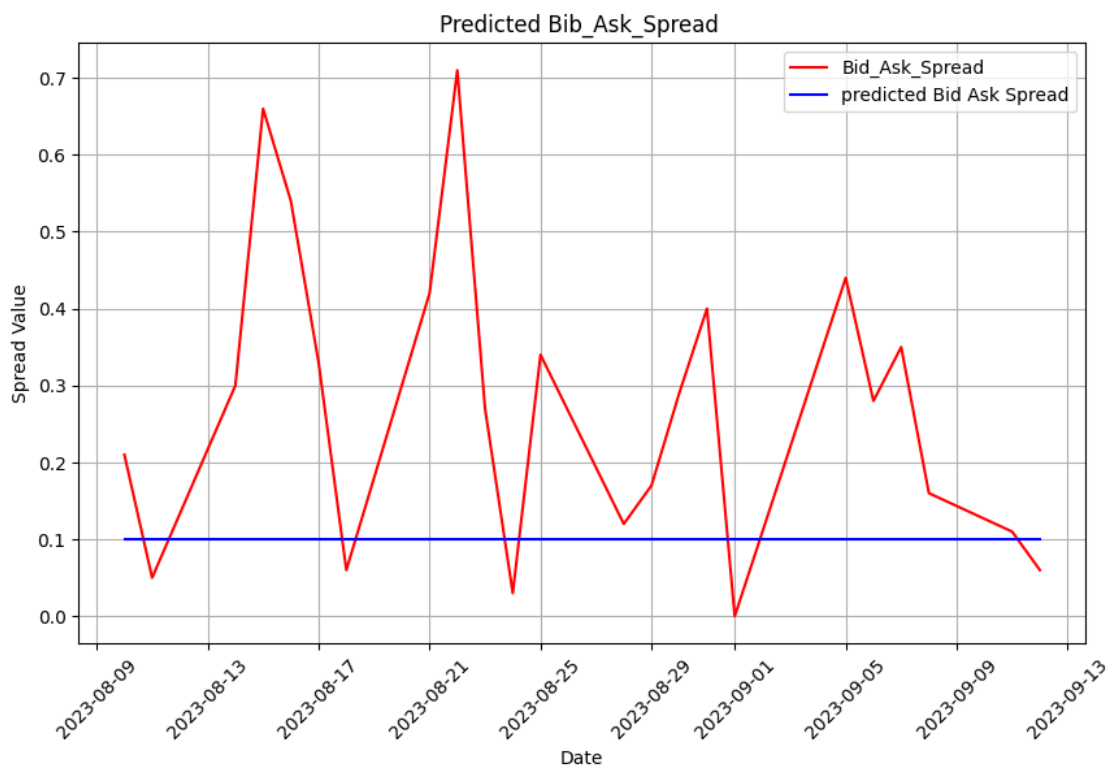
```
[77]: 0.10256099283656861
```

```
[78]: print(f'Mean Squared Error : {mse}')
```

```
Mean Squared Error : 0.10256099283656861
```

```
[79]: stock_data['Predicted Bid_Ask_Spread'] = 0.10
```

```
[81]: plt.figure(figsize=(10, 6))
plt.plot(stock_data.index, stock_data['Bid_Ask_Spread'],
         label='Bid_Ask_Spread', color='red',)
plt.plot(stock_data.index, stock_data['Predicted Bid_Ask_Spread'],
         label='predicted Bid Ask Spread', color='blue',)
plt.xlabel('Date')
plt.ylabel('Spread Value')
plt.title("Predicted Bib_Ask_Spread ")
plt.legend()
plt.grid(True)
plt.xticks(rotation=45)
plt.show()
```



```
[82]: model = BayesianRidge(compute_score=True)
model.fit(X_train, y_train)
```

```
[82]: BayesianRidge(compute_score=True)
```

```
[83]: model.coef_
```

```
[83]: array([0.00062784])
```

```
[85]: model.scores_
```

```
[85]: array([2.47296156, 5.27089326, 5.32294338, 5.32302901, 5.32289867])
```

```
[86]: print('Bayesian Ridge Regression Score :', model.score(X_test, y_test))
```

Bayesian Ridge Regression Score : -0.951045430618263

```
[87]: stock_data['Bayesian Ridge Regression Score'] = -0.95
```

```
[92]: plt.figure(figsize=(10, 6))
plt.plot(stock_data.index, stock_data['Bid_Ask_Spread'], label="↳
↳Bid_Ask_Spread", color='blue')
plt.plot(stock_data.index, stock_data['Bayesian Ridge Regression Score'],
↳label="Bayesian Ridge Regression Score", color='green')
plt.xlabel("Date")
plt.ylabel('Spread Value')
plt.title('Bayesian with Bid Ask Spread')
plt.legend()
plt.grid(True)
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

