

Gold Price Prediction

Ayu Siti Nasya N
MIP-ML-13



Dataset description

goldstock.csv

dataset of daily gold prices spanning from January 19, 2014, to January 22, 2024, obtained from Nasdaq




```

RangeIndex: 2511 entries, 0 to 2510
Data columns (total 6 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Date        2511 non-null   object
1   Close       2511 non-null   float64
2   Volume      2511 non-null   float64
3   Open        2511 non-null   float64
4   High        2511 non-null   float64
5   Low         2511 non-null   float64
dtypes: float64(5), object(1)

```

- 1.Date: A unique identifier for each trading day
- 2.Close: Closing price of gold on the respective date
- 3.Volume: Gold trading volume on the corresponding date
- 4.Open: Opening price of gold on the respective date
- 5.High: The highest recorded price of gold during the trading day
- 6.Low: The lowest price recorded for gold in the trading day.

Exploratory Data Analysis

01

Visualization of distribution

02

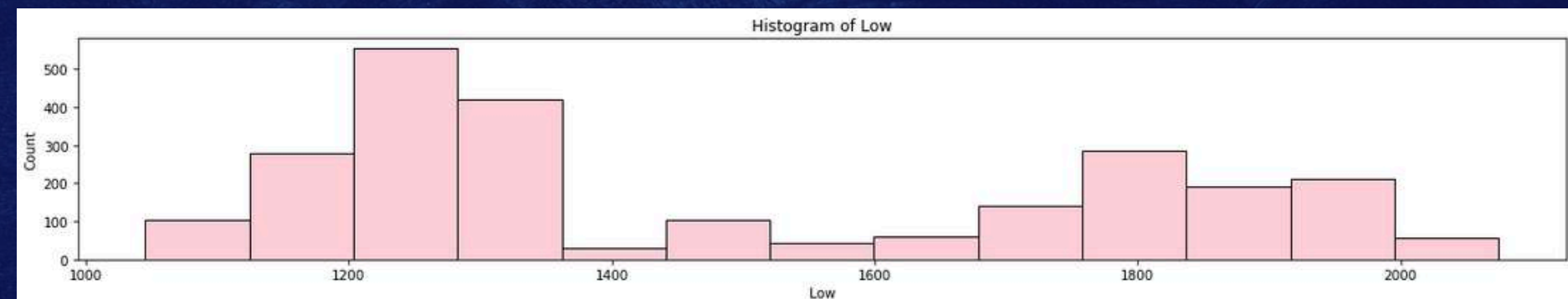
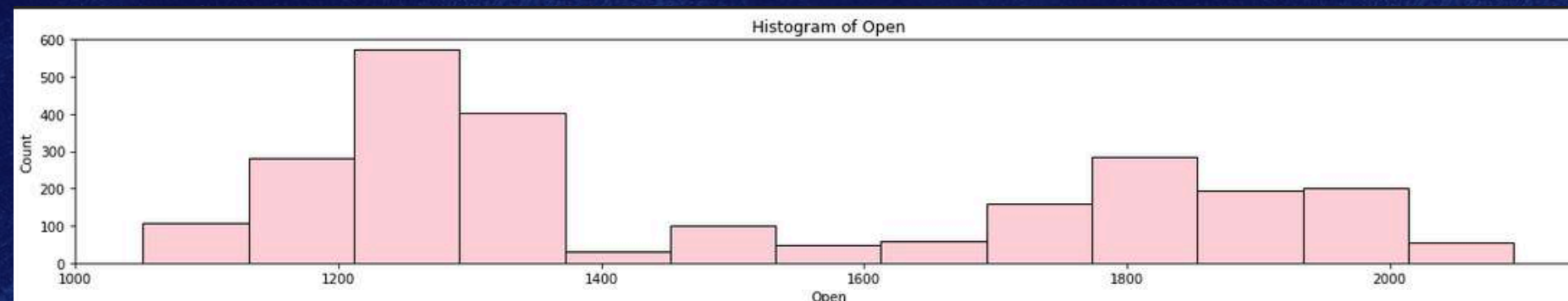
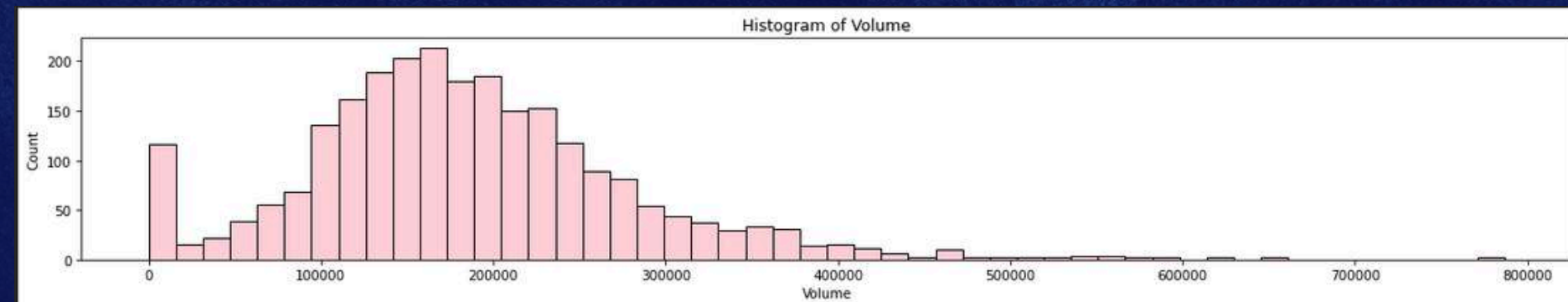
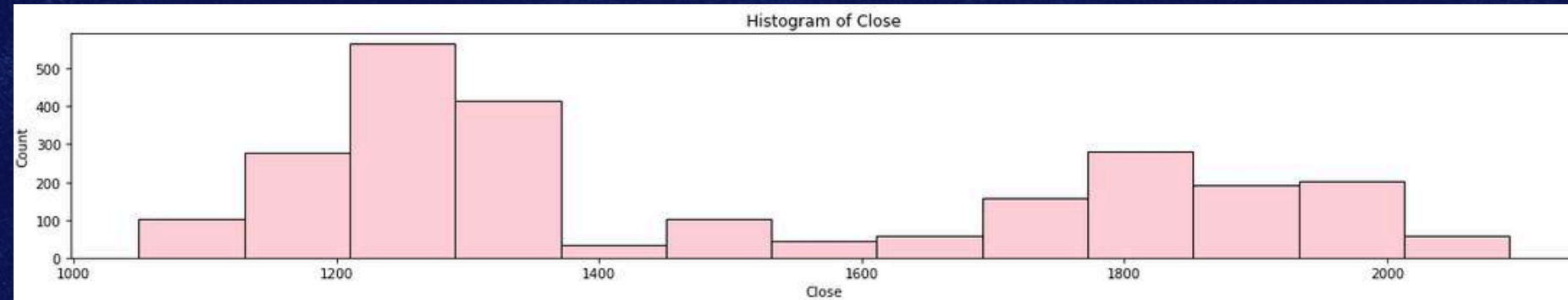
Analysis of gold price over time

03

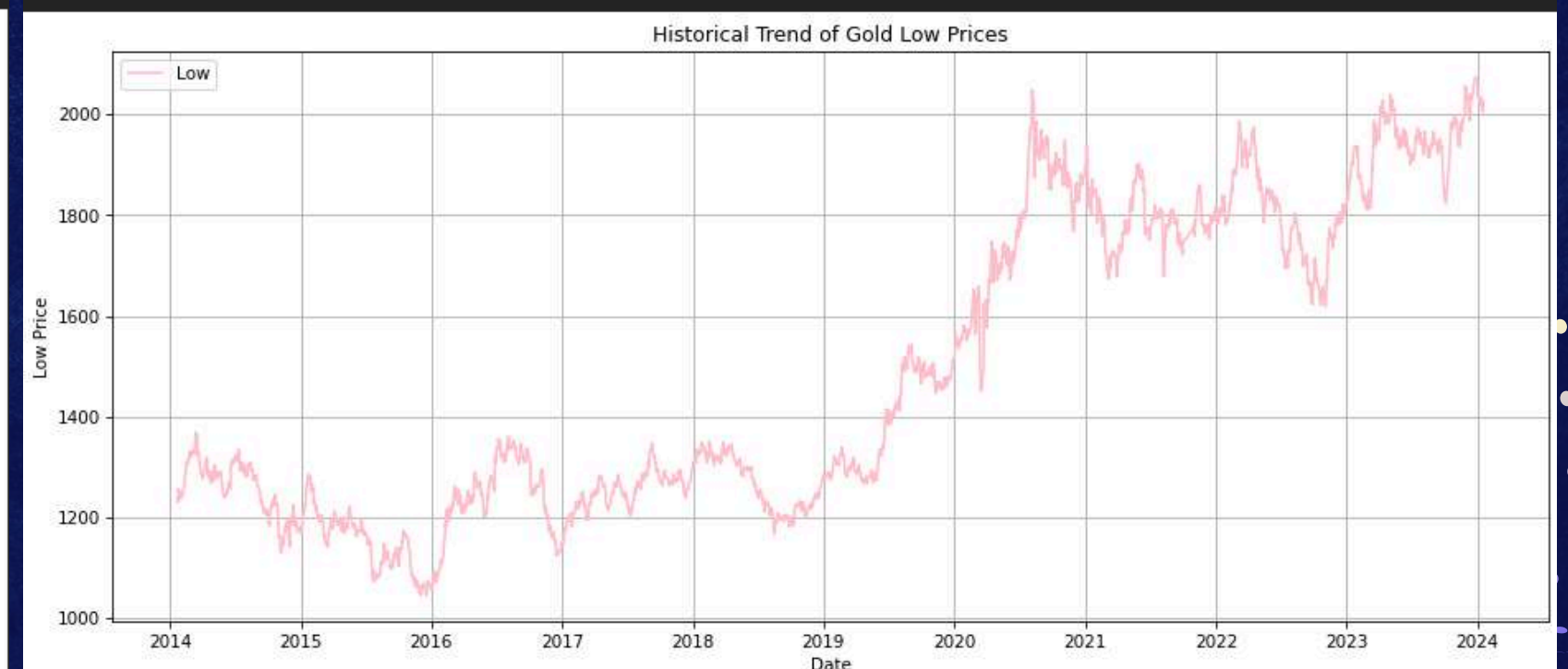
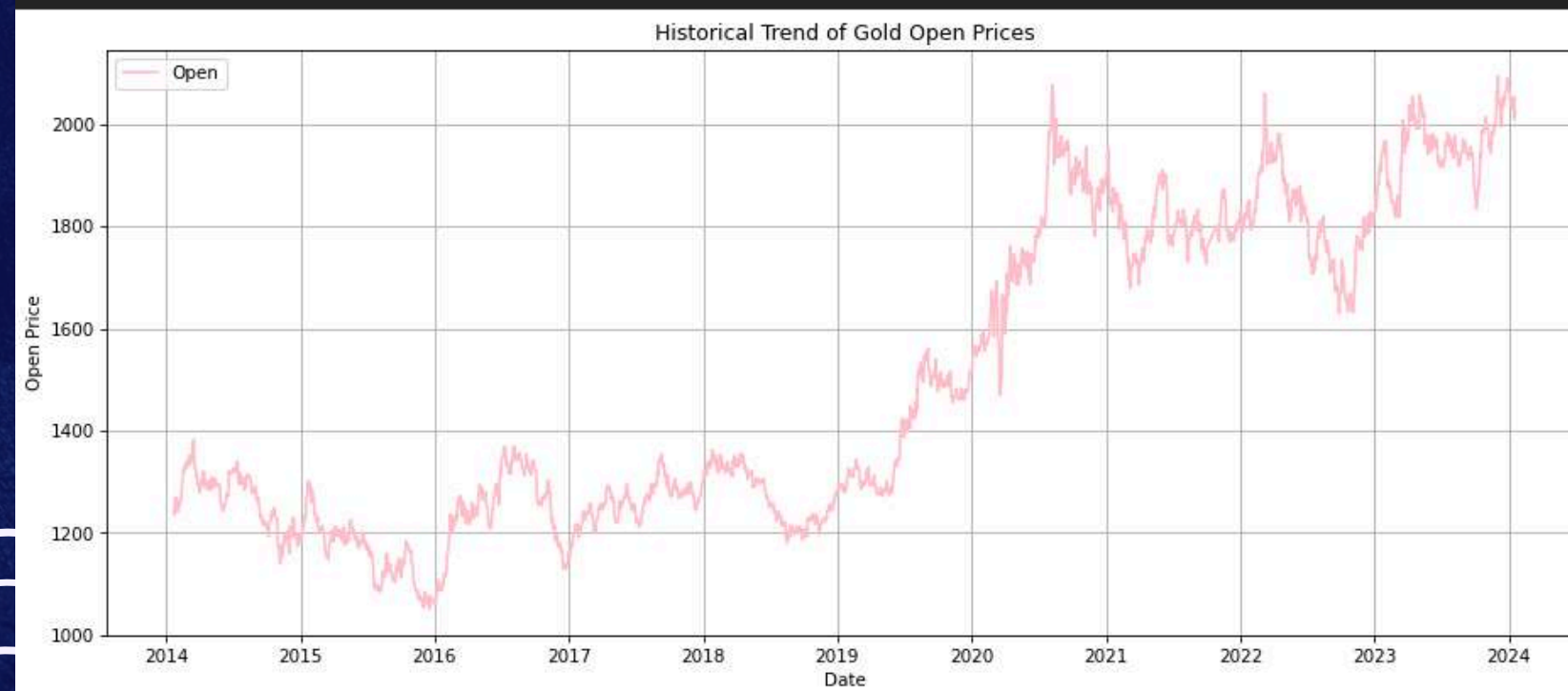
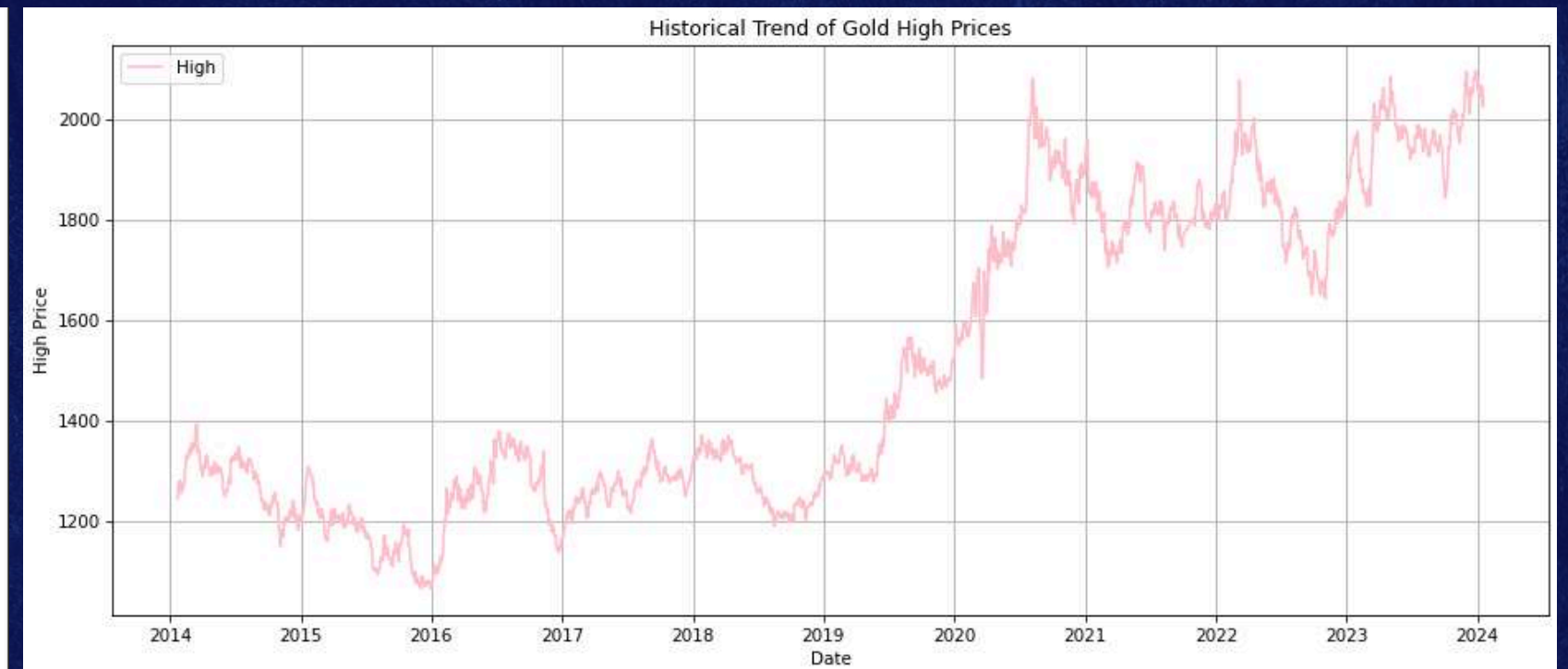
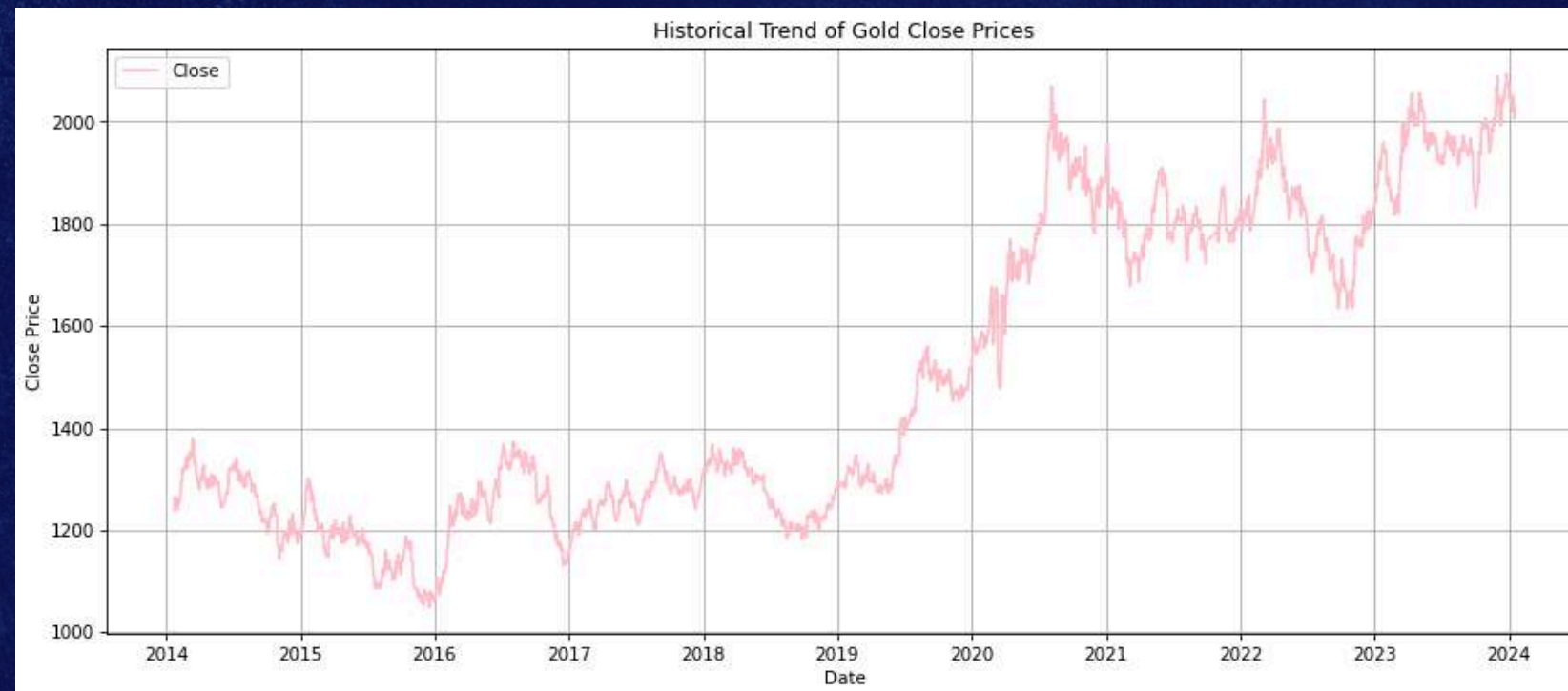
Time series analysis

01

Visualization of Distribution



Analysis of Gold Prices Over Time

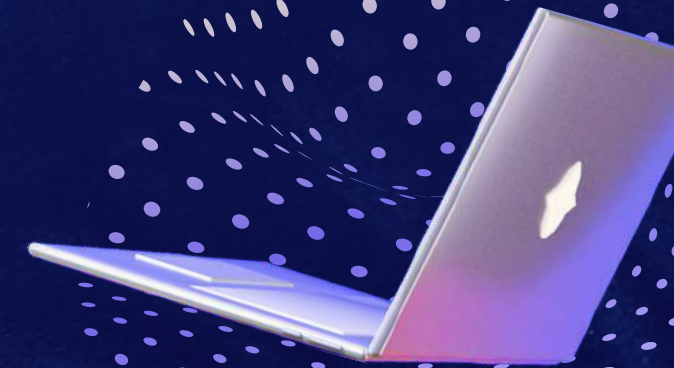


Analysis of Gold Prices Over Time

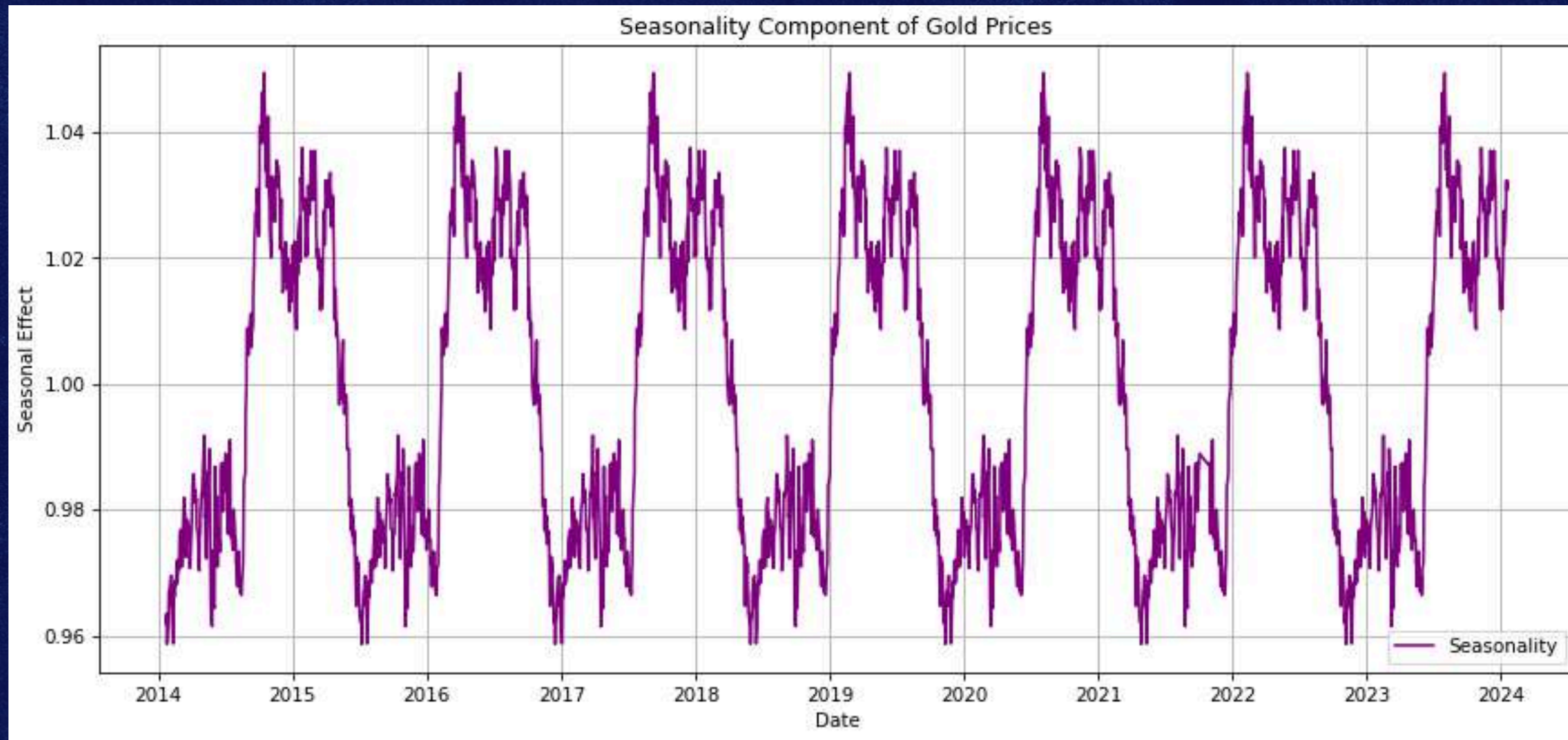


02

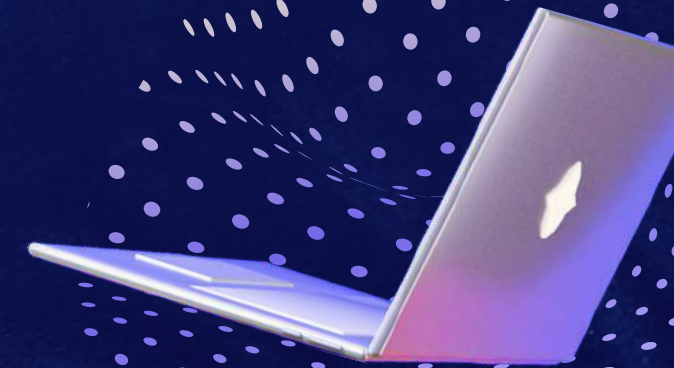
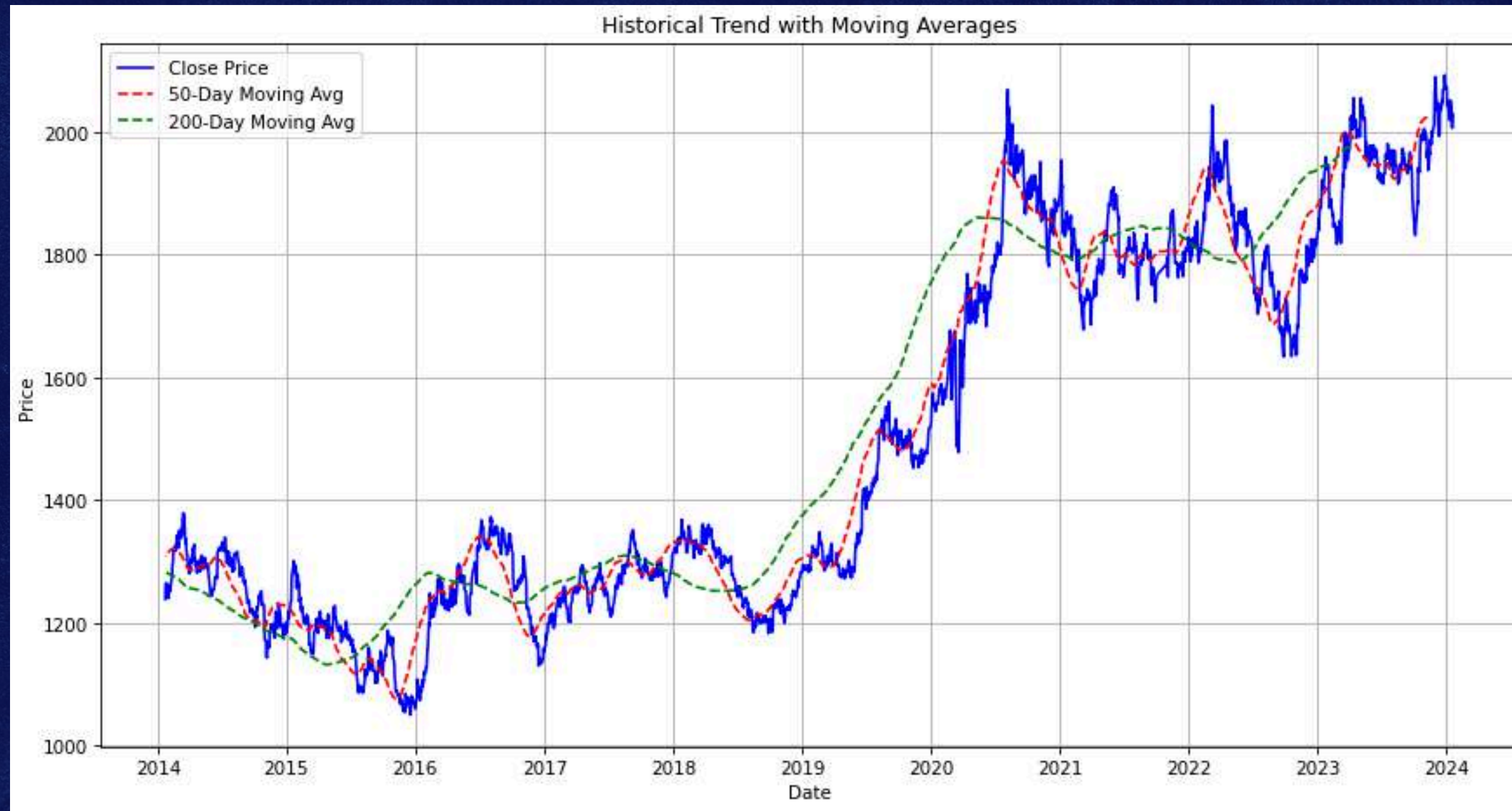
Analysis of Gold Prices Over Time



Time Series Analysis



Time Series Analysis



Gold Price Forecasting

01

ARIMA

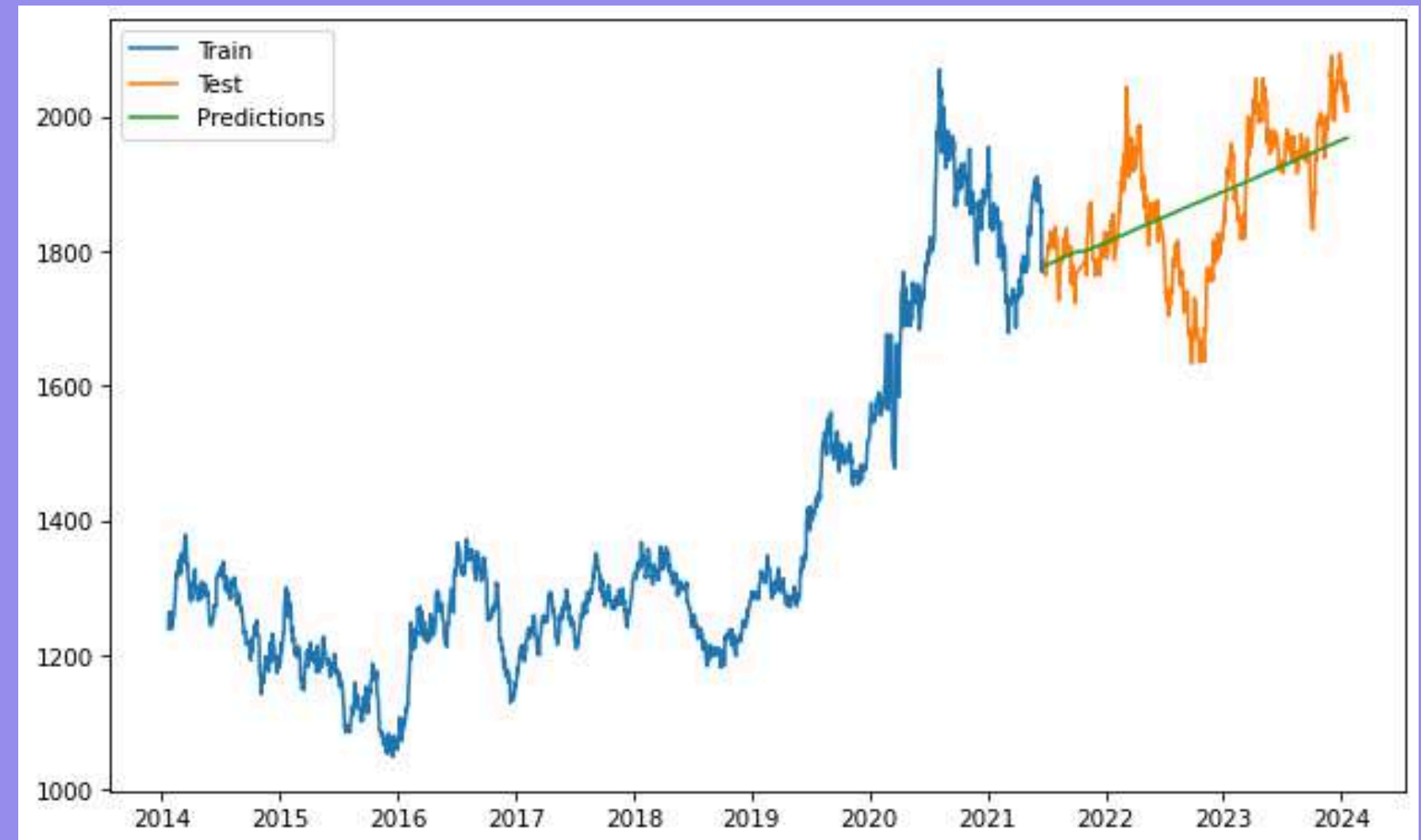
02

Random Forest Regression

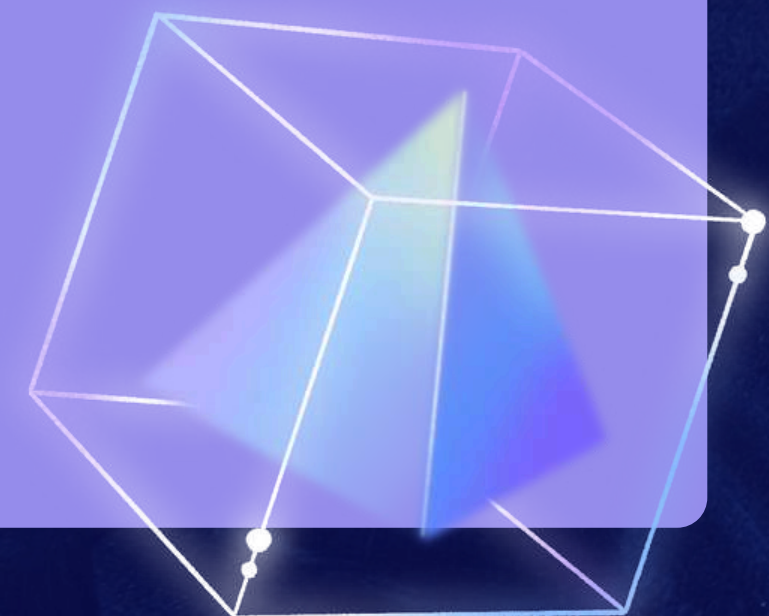
01

ARIMA MODEL

```
SARIMAX Results
=====
Dep. Variable:      Close      No. Observations:      1869
Model:              ARIMA(0, 2, 1)  Log Likelihood      -7549.167
Date:              Tue, 25 Jun 2024  AIC              15102.334
Time:              11:43:53    BIC              15113.398
Sample:            0          HQIC              15106.410
                  - 1869
Covariance Type:    opg
=====
              coef    std err          z      P>|z|      [0.025    0.975]
-----
ma.L1         -0.9996      0.003    -314.053      0.000      -1.006     -0.993
sigma2         189.6777      3.101      61.167      0.000      183.600     195.755
=====
Ljung-Box (L1) (Q):                1.70    Jarque-Bera (JB):                4833.96
Prob(Q):                           0.19    Prob(JB):                     0.00
Heteroskedasticity (H):              2.53    Skew:                         -0.40
Prob(H) (two-sided):                0.00    Kurtosis:                     10.84
=====
```



$d = 0$
 $p = 2$
 $q = 1$



01

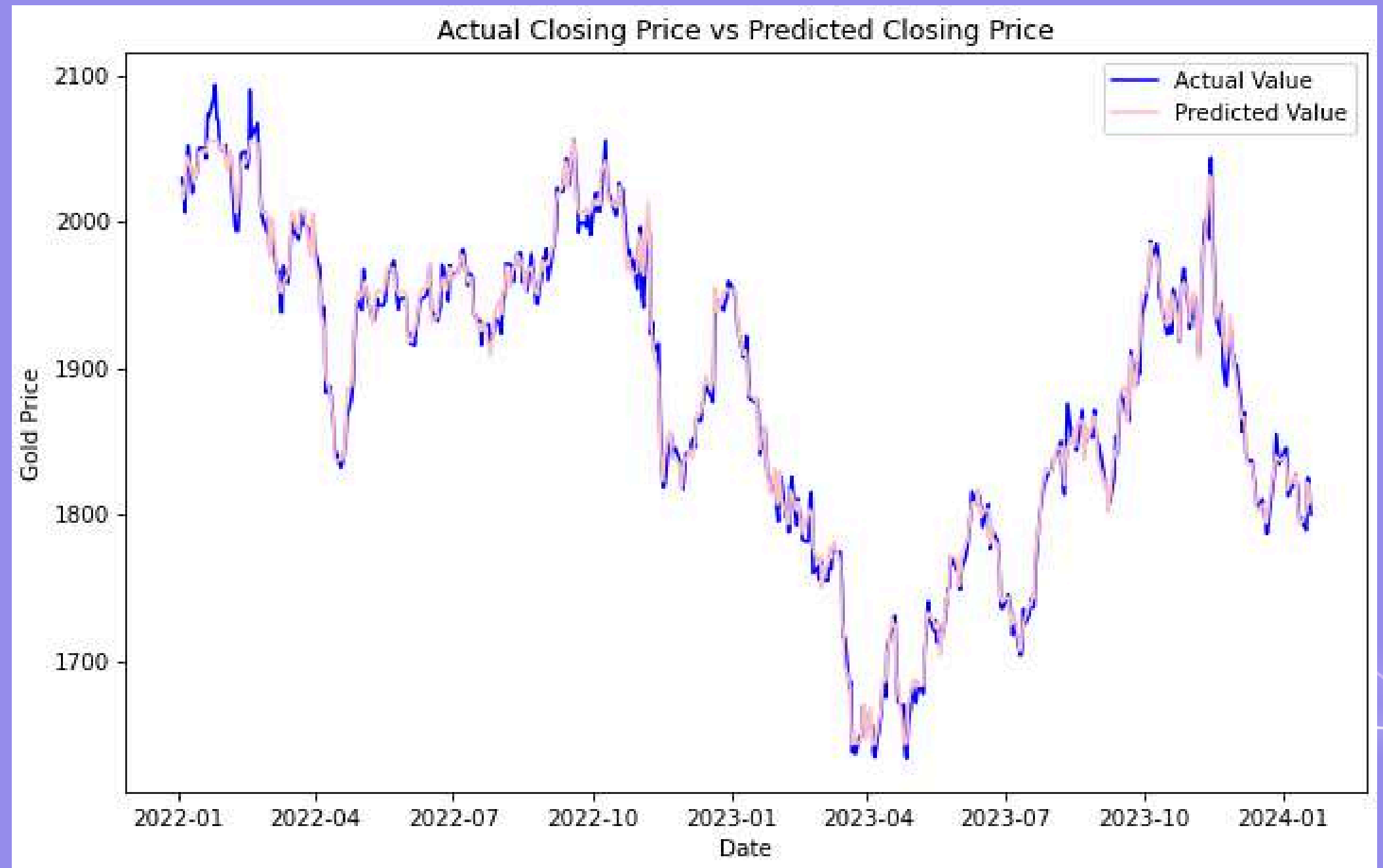
RANDOM FOREST REGRESSION

'Parameter :

- criterion: 'squared_error'
- 'max_depth': 9
- n_estimators: 100

R squared error :

0.9919293591886352

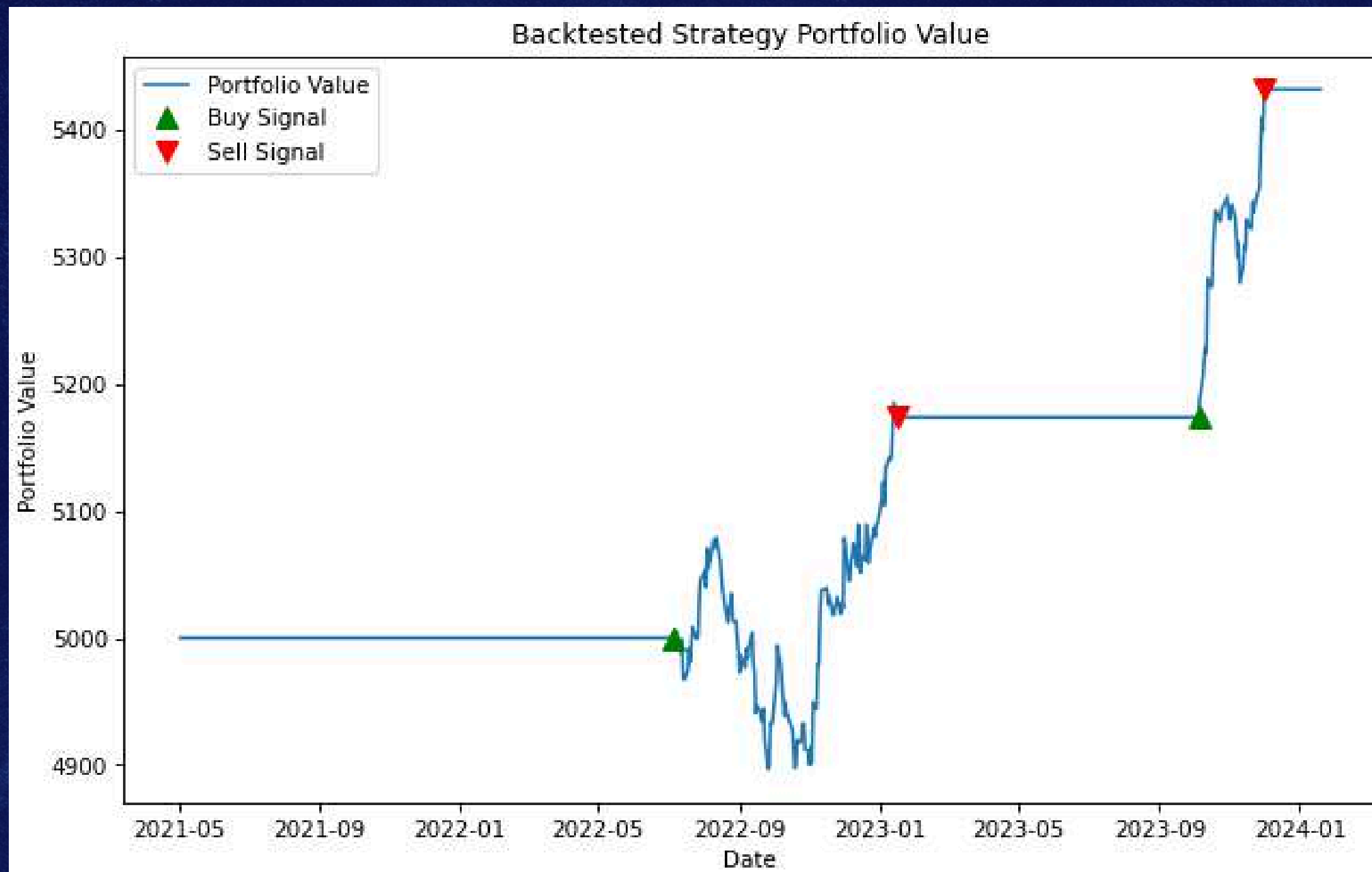




Backtesting

Moving average crossover
strategy





- Number of trades: 2.0
- Initial capital: \$5,000.00
- Final capital: \$5,431.30
- Total return: 8.63%

Market Analysis

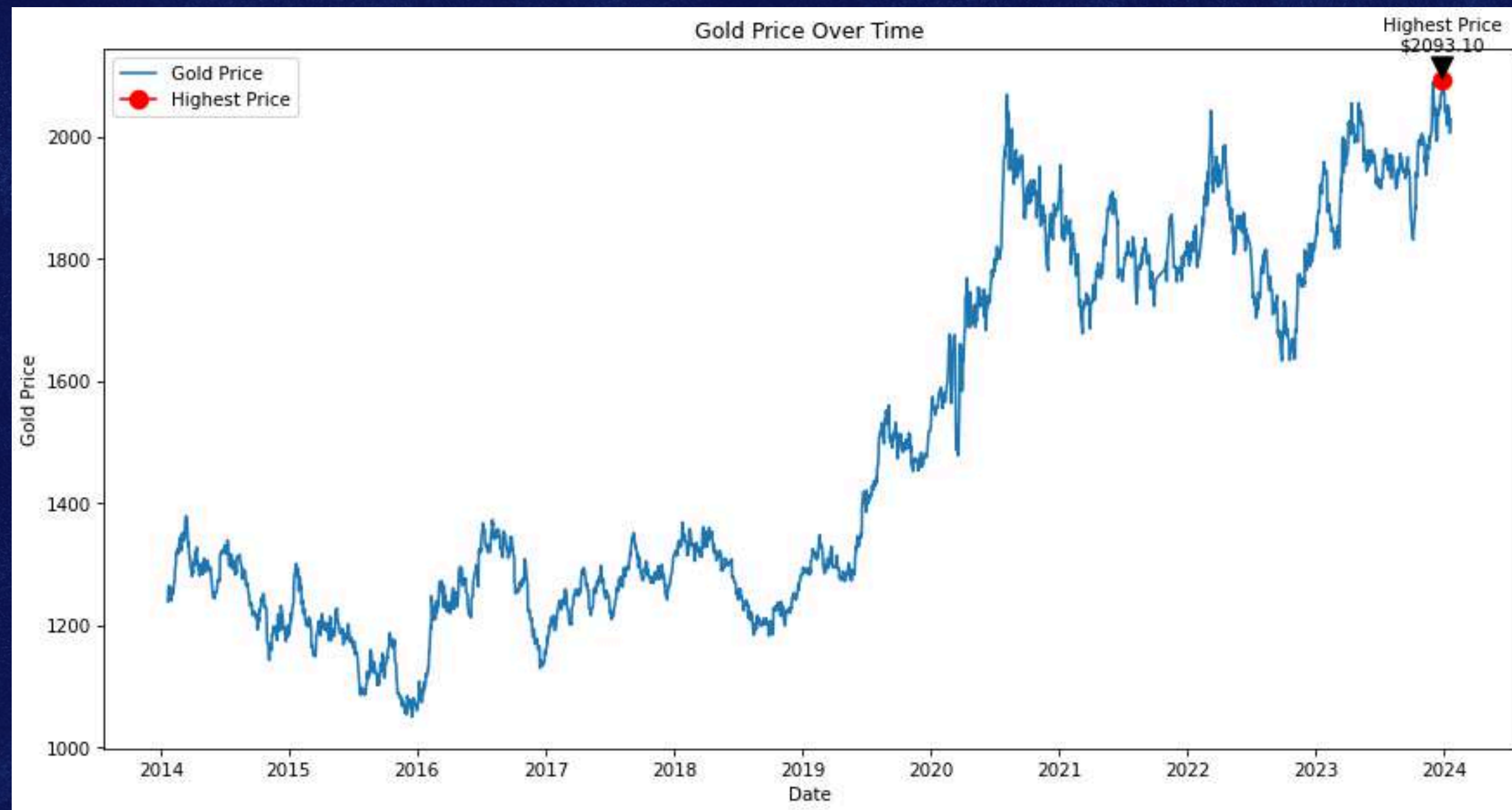
01

Highest price of gold

02

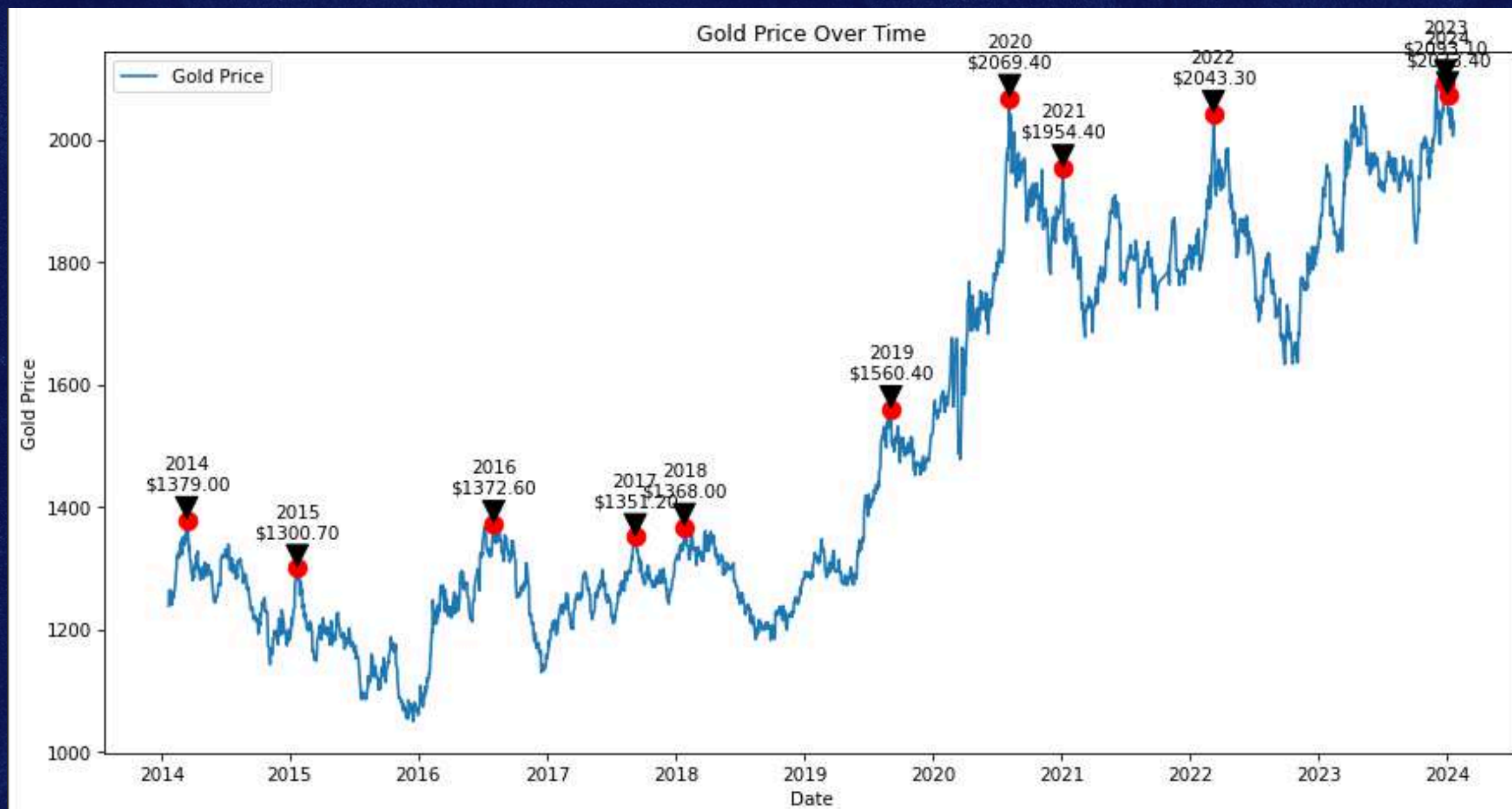
Most drastic increase in gold
prices

Highest Gold Price

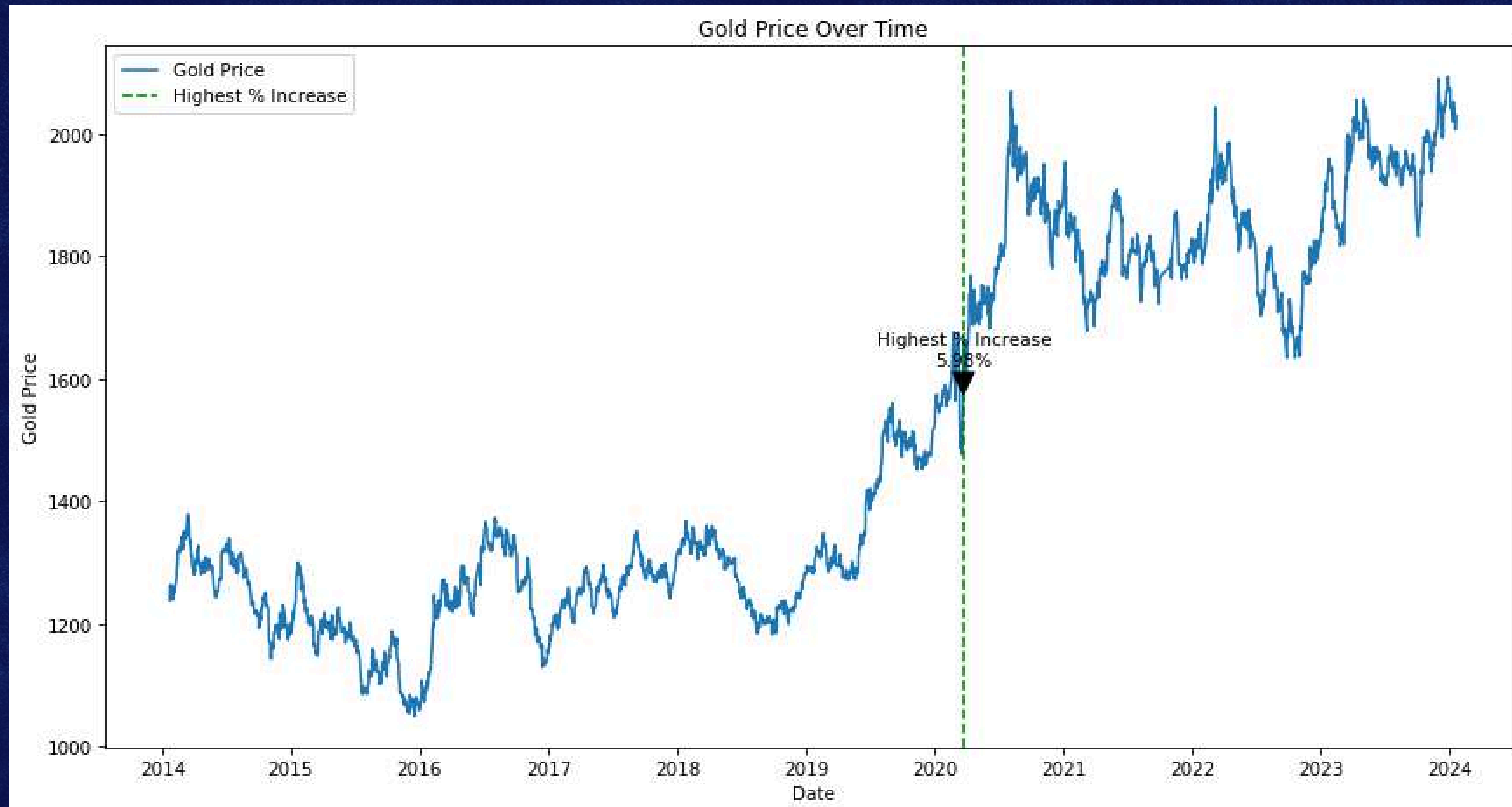


Highest gold price: \$2093.1 on 2023-12-27

Highest Gold Price each Year



Most Drastic Increase



Statistical Analysis



Statistic Descriptive

	Close	Volume	Open	High	Low	Year	Pct Change
count	2493.000000	2493.000000	2493.000000	2493.000000	2493.000000	2493.000000	2492.000000
mean	1496.354272	186034.698355	1496.342479	1506.056839	1486.518091	2018.525872	0.000242
std	298.575898	97739.426777	298.860270	301.009257	296.172036	2.883059	0.009324
min	1049.600000	1.000000	1051.500000	1062.700000	1045.400000	2014.000000	-0.049854
25%	1248.600000	126411.000000	1248.500000	1256.600000	1241.700000	2016.000000	-0.004395
50%	1331.400000	175222.000000	1332.000000	1340.900000	1324.500000	2019.000000	0.000208
75%	1803.800000	234908.000000	1803.800000	1813.900000	1791.800000	2021.000000	0.005051
max	2093.100000	787217.000000	2094.400000	2098.200000	2074.600000	2024.000000	0.059771

Correlation Heatmap



**Thank
you very
much!**

