STOCK MARKET ANALYSIS

Date	Open	High	Low	Close	Adj Close	Volume	company_name
2024-05-28	179.929993	182.240005	179.490005	182.149994	182.149994	29927000	AMAZON
2024-05-29	181.699997	184.080002	181.550003	182.020004	182.020004	32009300	AMAZON
2024-05-30	181.309998	181.339996	178.360001	179.320007	179.320007	29249200	AMAZON
2024-05-31	178.300003	179.210007	173.869995	176.440002	176.440002	58903900	AMAZON
2024-06-03	177.699997	178.699997	175.919998	178.339996	178.339996	30786600	AMAZON
2024-06-04	177.639999	179.820007	176.440002	179.339996	179.339996	27198400	AMAZON

Import Libraries: Import necessary libraries.

Set Styles: Set styles for plotting.

Read Stock Data: Set up the stock symbols and download their data from Yahoo Finance.

Add Company Names: Add a column to each DataFrame indicating the company name.

Concatenate Data: Combine all individual DataFrames into one.

Using `globals()` is a sloppy way of setting the `DataFrame` names, but it's simple. Now we have our data, let's perform some basic data analysis and check our data.

	Open	High	Low	Close	Adj Close	Volume
count	251.000000	251.000000	251.000000	251.000000	251.000000	2.510000e+02
mean	182.746534	184.211999	181.367410	182.821753	182.296098	5.750452e+07
std	8.738375	8.552274	8.749900	8.690825	8.641298	1.844647e+07
min	165.350006	166.399994	164.080002	165.000000	164.776505	1.421504e+07
25%	175.244995	177.025002	173.660004	175.280006	174.800789	4.657160e+07
50%	183.419998	184.899994	181.809998	183.630005	182.988205	5.302030e+07
75%	190.239998	191.540001	189.189995	190.139999	189.854996	6.438910e+07
max	198.020004	199.619995	197.000000	198.110001	197.589523	1.632241e+08

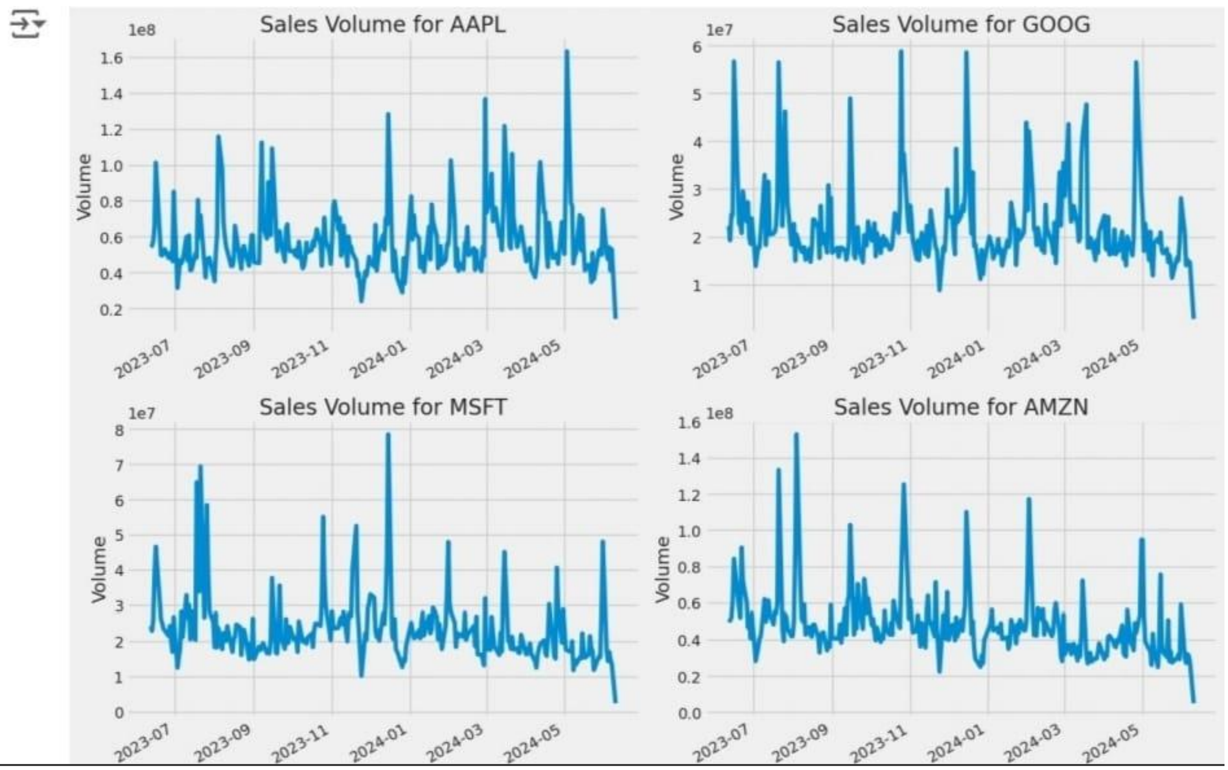
We have only 255 records in one year because weekends are not included in the data.

Closing Price

The closing price is the last price at which the stock is traded during the regular trading day. A stock’s closing price is the standard benchmark used by investors to track its performance over time.

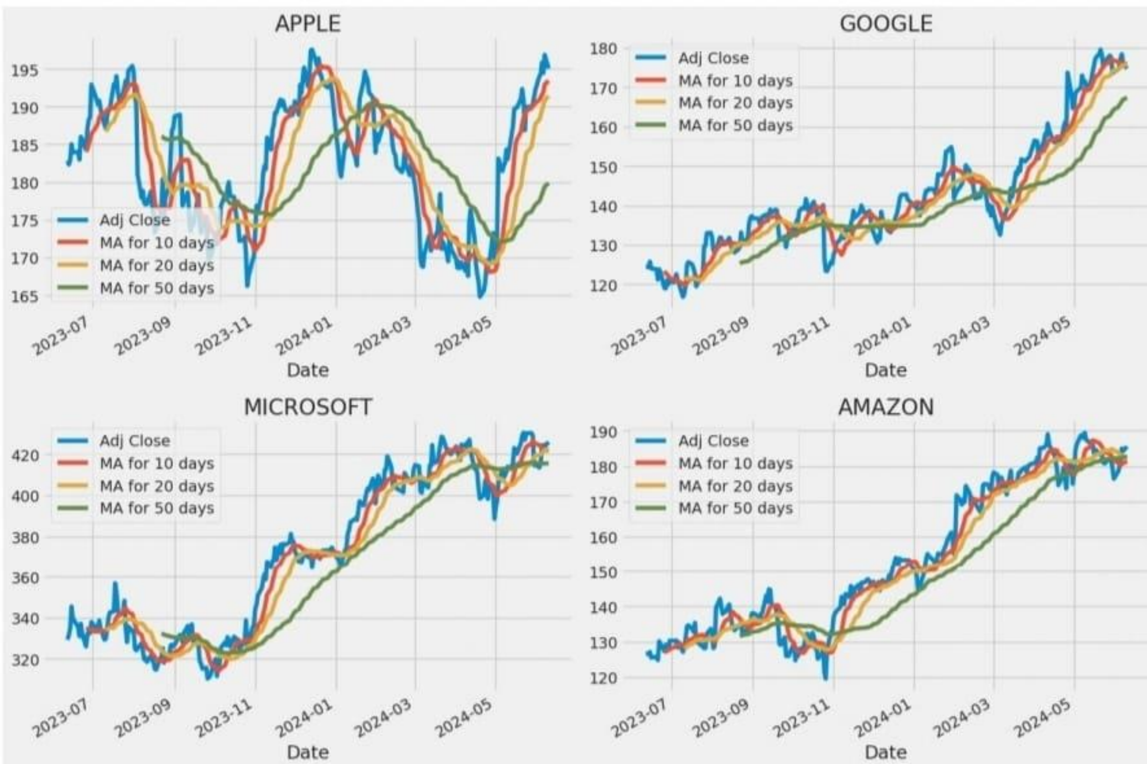
Volume of Sales

Volume is the amount of an asset or security that changes hands over some period of time, often over the course of a day. For instance, the stock trading volume would refer to the number of shares of security traded between its daily open and close. Trading volume, and changes to volume over the course of time, are important inputs for technical traders.



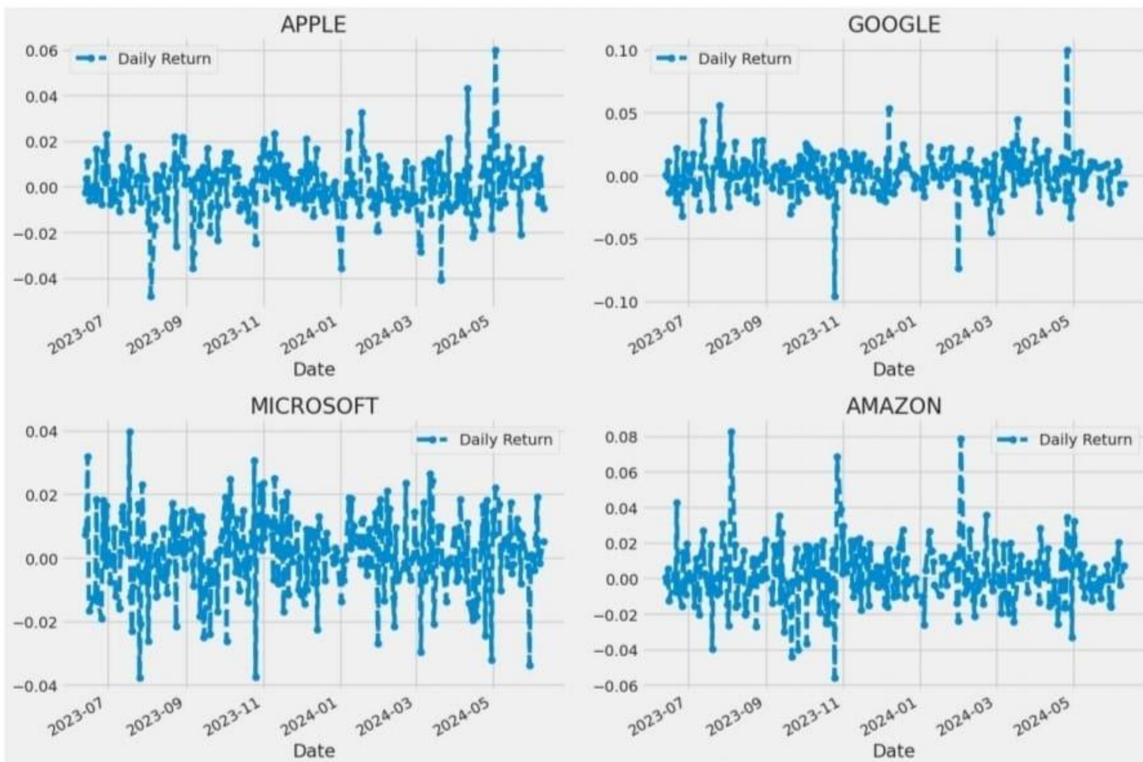
What was the moving average of the various stocks?

The moving average (MA) is a simple technical analysis tool that smooths out price data by creating a constantly updated average price. The average is taken over a specific period of time, like 10 days, 20 minutes, 30 weeks, or any time period the trader chooses.

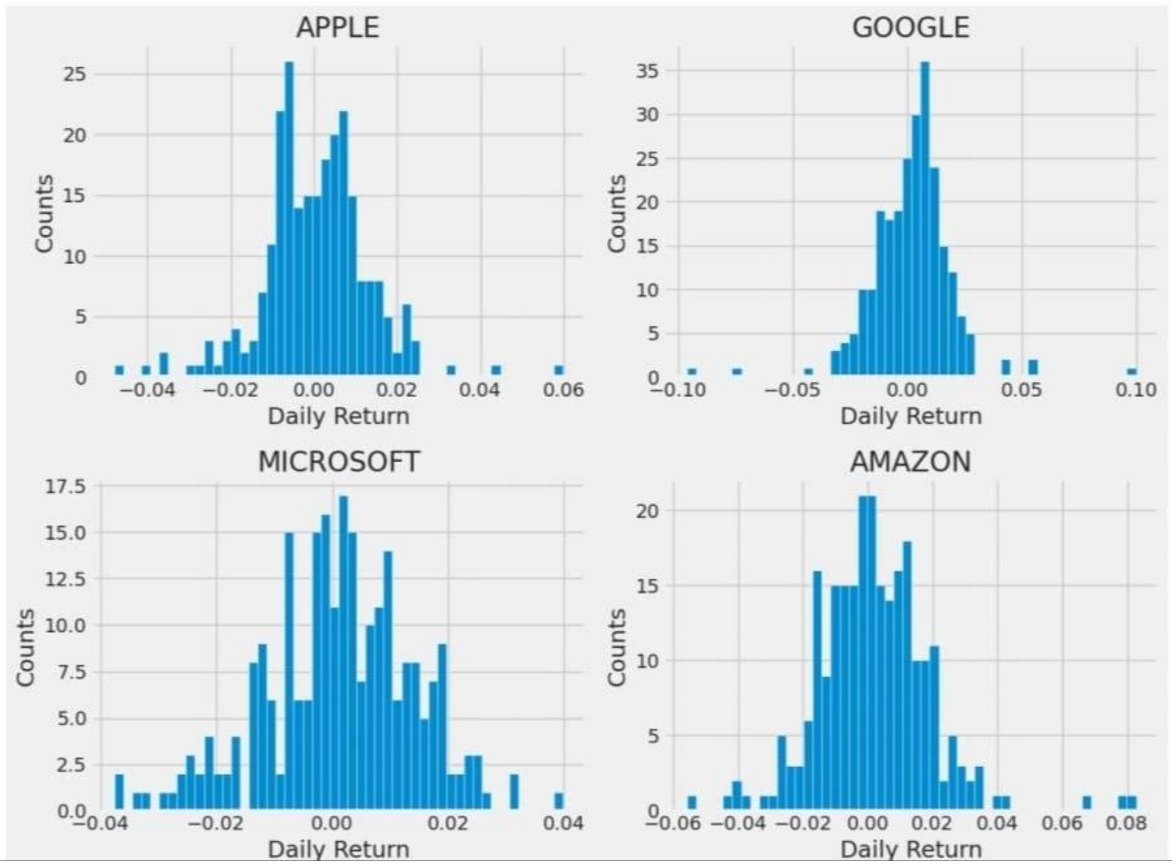


What was the daily return of the stock on average?

Now that we've done some baseline analysis, let's go ahead and dive a little deeper. We're now going to analyze the risk of the stock. In order to do so we'll need to take a closer look at the daily changes of the stock, and not just its absolute value. Let's go ahead and use pandas to retrieve teh daily returns for the Apple stock.



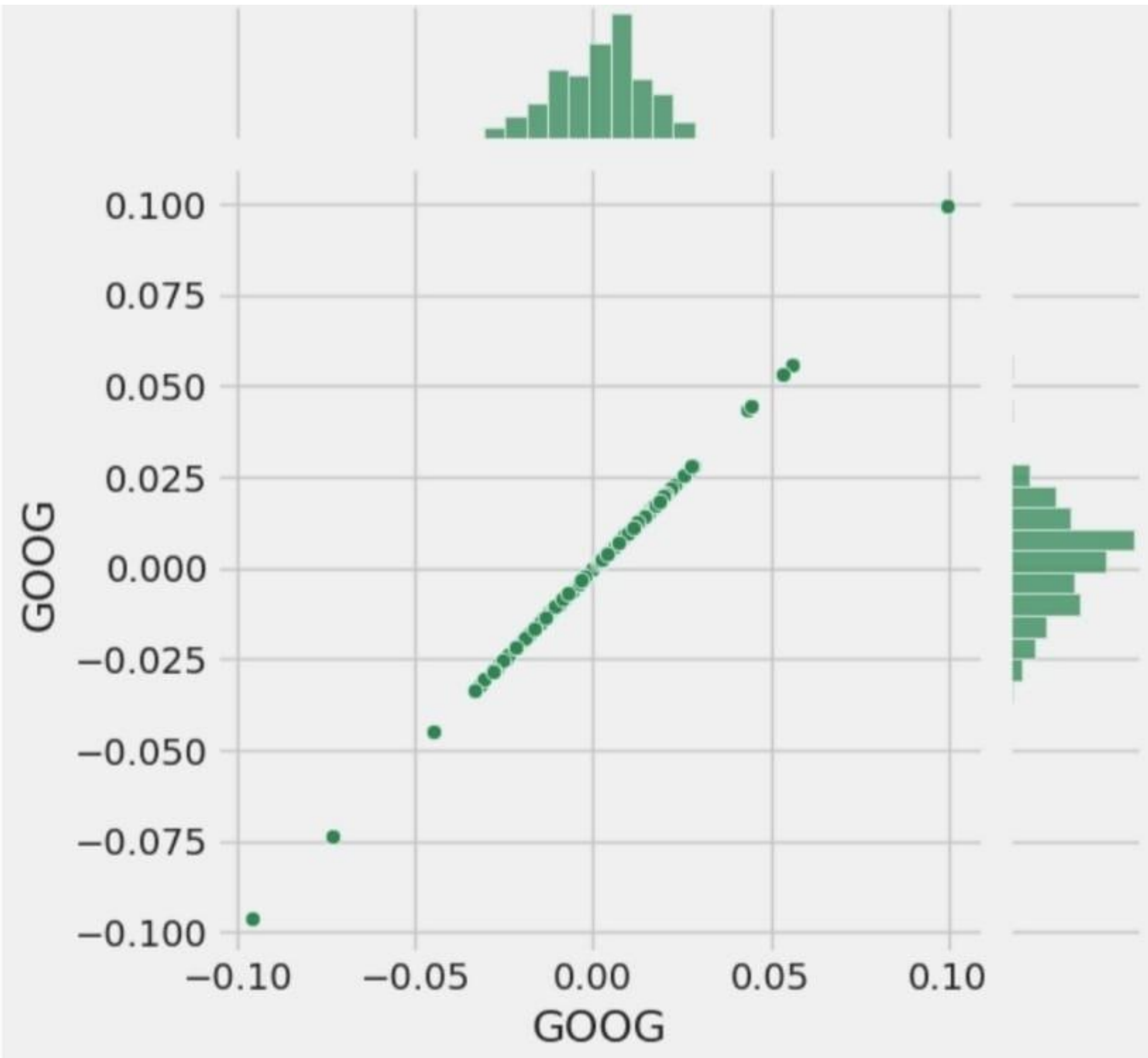
Great, now let's get an overall look at the average daily return using a histogram. We'll use seaborn to create both a histogram and kde plot on the same figure.

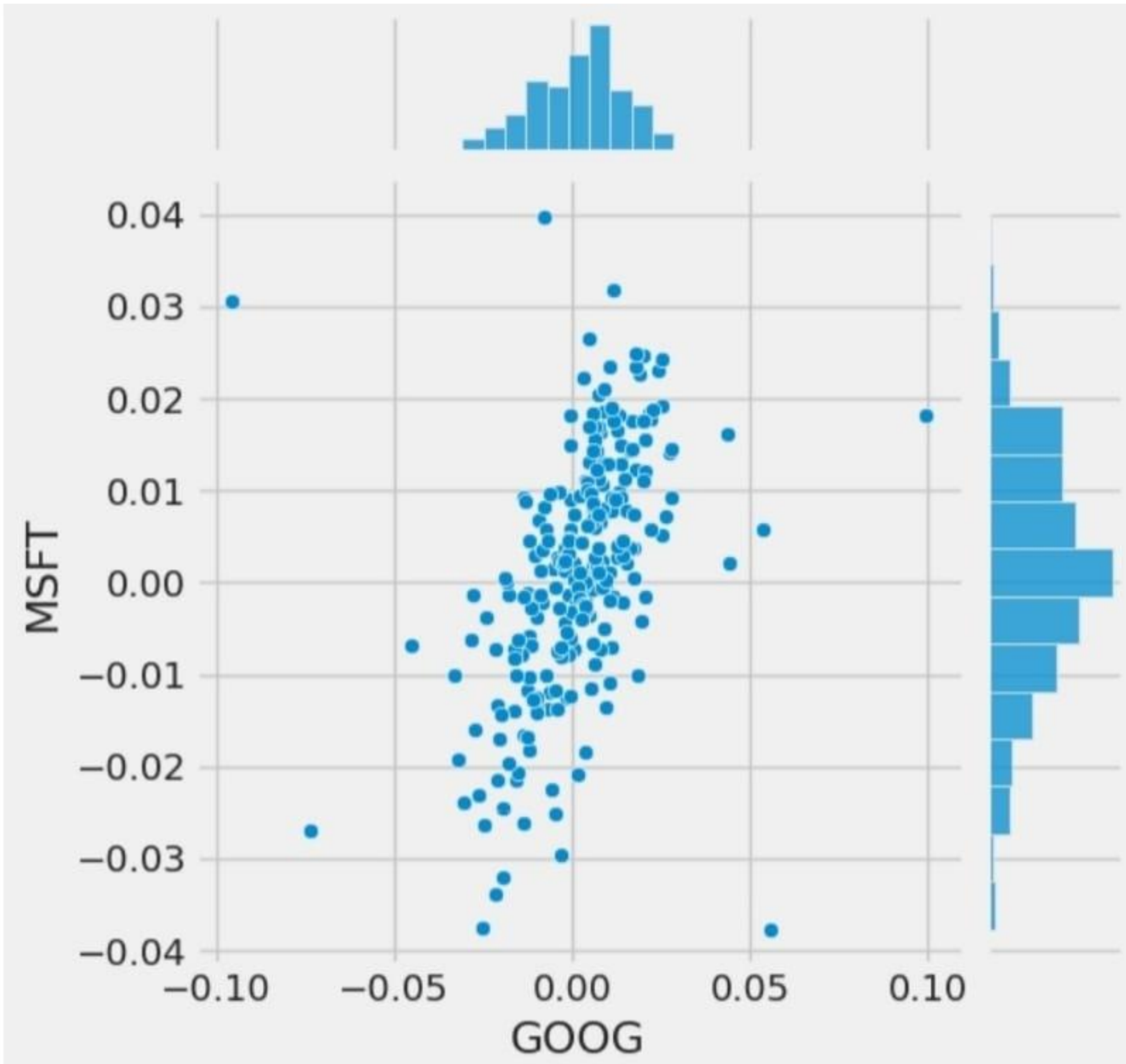


What was the correlation between different stocks closing prices?

Correlation is a statistic that measures the degree to which two variables move in relation to each other which has a value that must fall between -1.0 and +1.0. Correlation measures association, but doesn't show if x causes y or vice versa — or if the association is caused by a third factor[1].

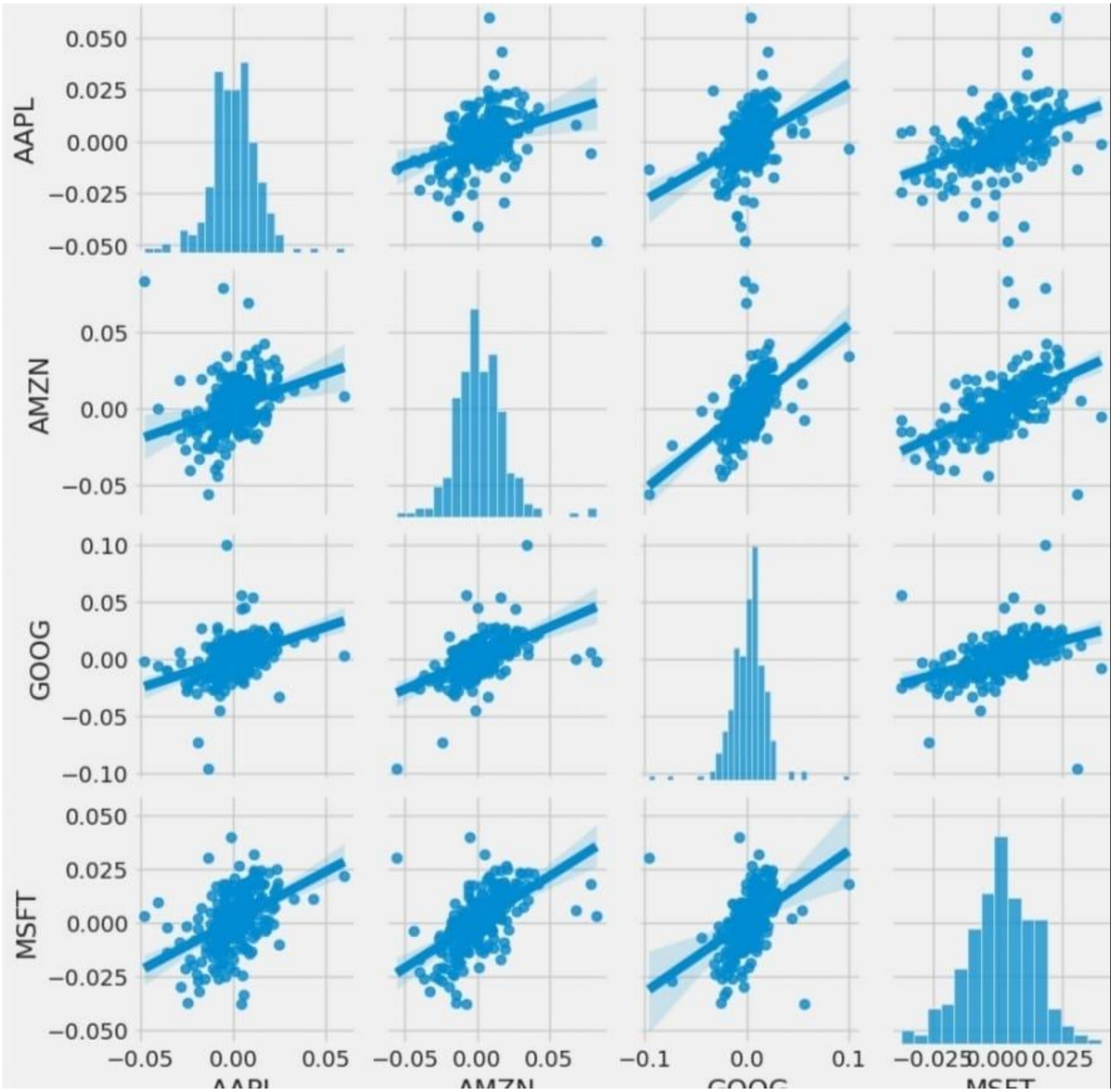
Now what if we wanted to analyze the returns of all the stocks in our list? Let's go ahead and build a DataFrame with all the ['Close'] columns for each of the stocks dataframes.





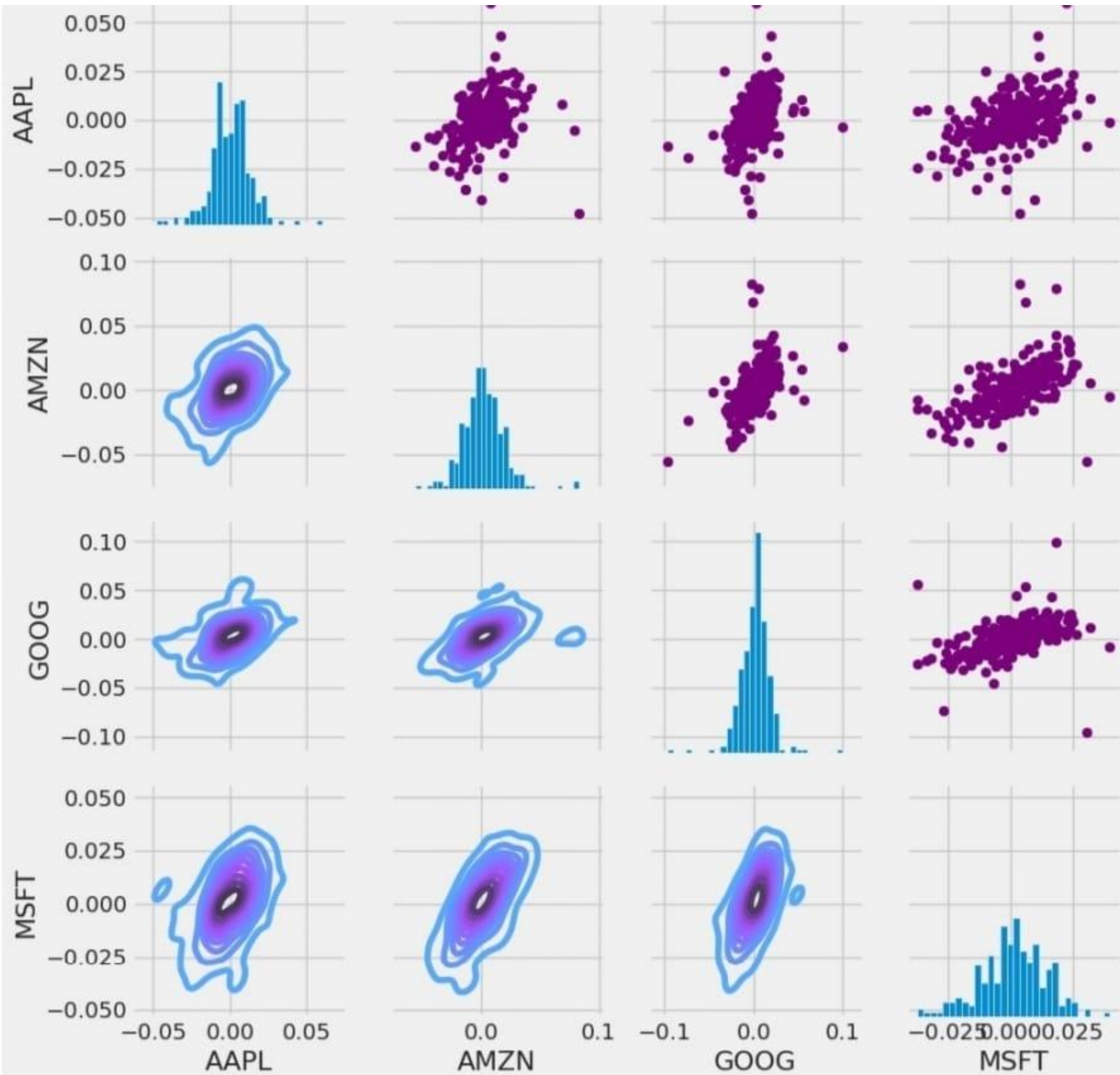
So now we can see that if two stocks are perfectly (and positivley) correlated with each other a linear relationship bewteen its daily return values should occur.

Seaborn and pandas make it very easy to repeat this comparison analysis for every possible combination of stocks in our technology stock ticker list. We can use sns.pairplot() to automatically create this plot

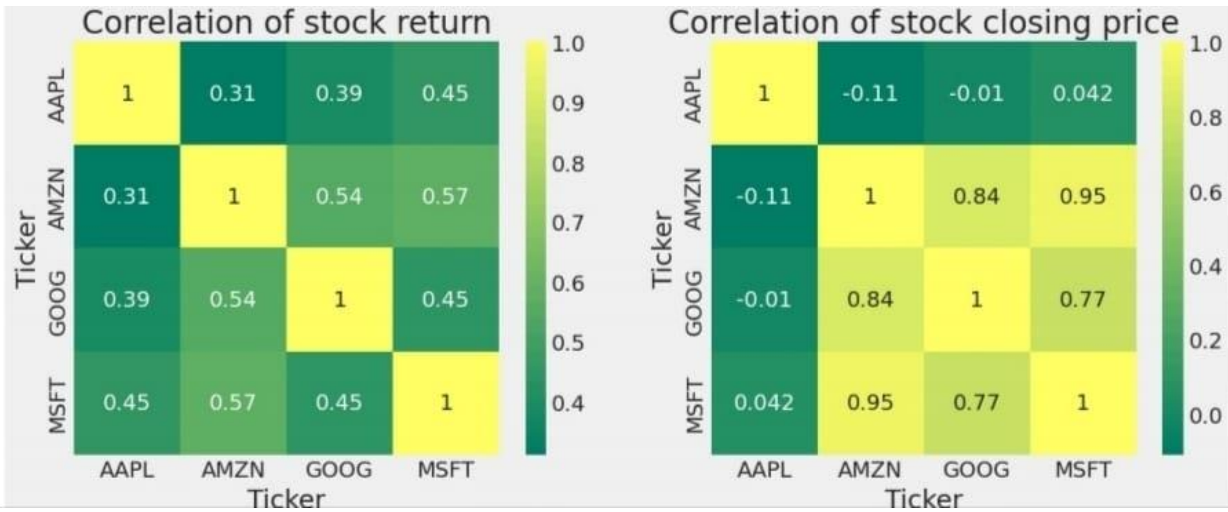


Above we can see all the relationships on daily returns between all the stocks. A quick glance shows an interesting correlation between Google and Amazon daily returns. It might be interesting to investigate that individual comaprison.

While the simplicity of just calling sns.pairplot() is fantastic we can also use sns.PairGrid() for full control of the figure, including what kind of plots go in the diagonal, the upper triangle, and the lower triangle. Below is an example of utilizing the full power of seaborn to achieve this result.



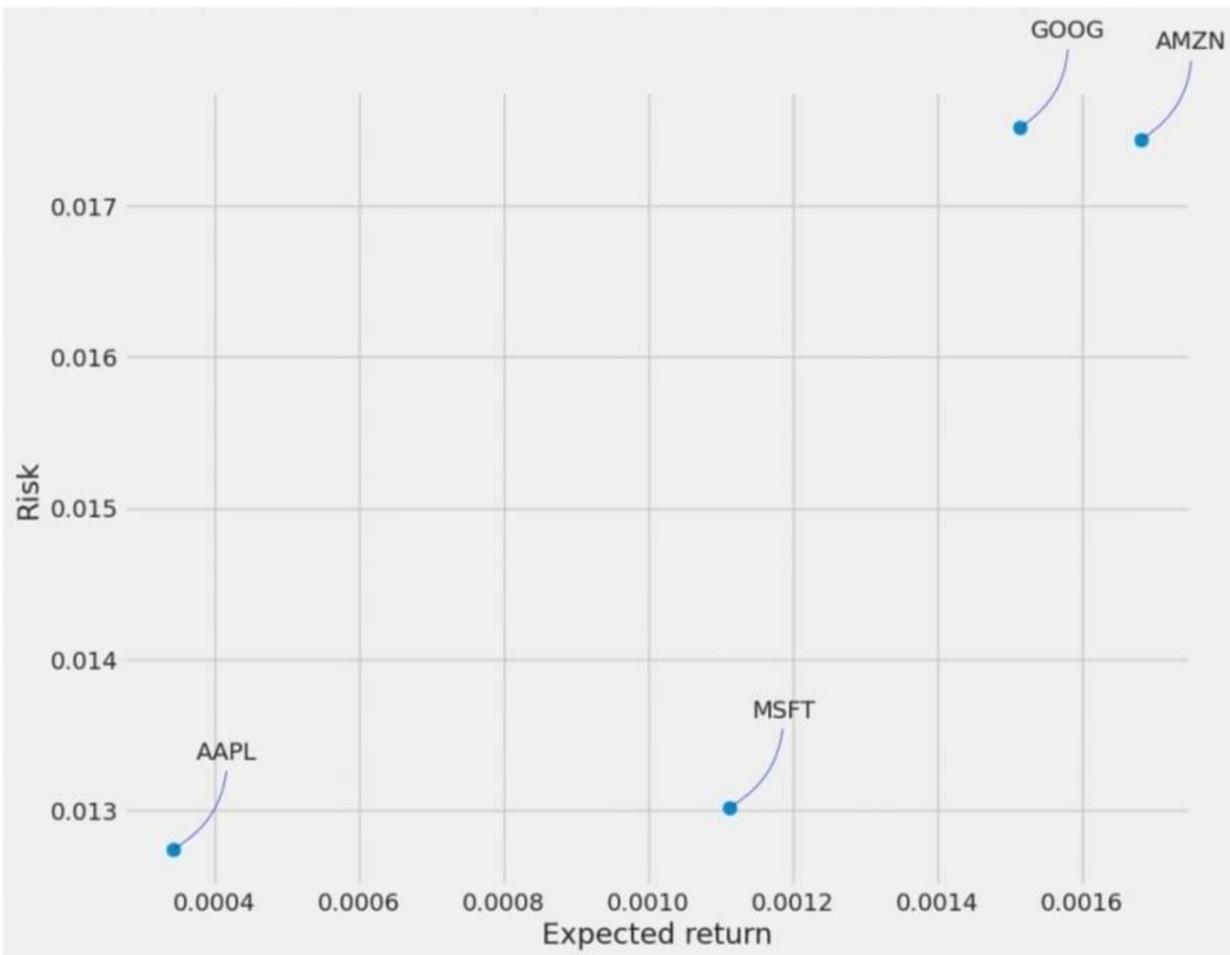
Finally, we could also do a correlation plot, to get actual numerical values for the correlation between the stocks' daily return values. By comparing the closing prices, we see an interesting relationship between Microsoft and Apple.



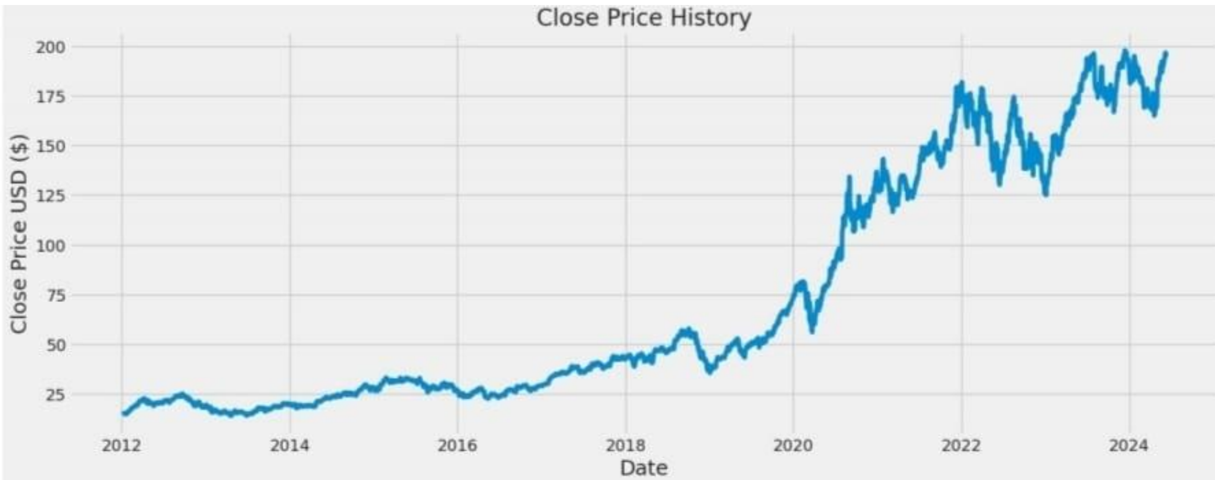
Just like we suspected in our PairPlot we see here numerically and visually that Microsoft and Amazon had the strongest correlation of daily stock return. It's also interesting to see that all the technology comapnies are positively correlated.

How much value do we put at risk by investing in a particular stock?

There are many ways we can quantify risk, one of the most basic ways using the information we've gathered on daily percentage returns is by comparing the expected return with the standard deviation of the daily returns.



Predicting the closing price stock price of APPLE inc



ARIMA models the level of the time series and captures trends and seasonality.

GARCH models the volatility (variance) of the returns series, capturing periods of high and low volatility.

```
ARIMA model summary for AMZN:

                               SARIMAX Results
==============================================================================
Dep. Variable:                  Close   No. Observations:                 3129
Model:                 ARIMA(5, 1, 0)   Log Likelihood               -5962.806
Date:                Mon, 10 Jun 2024   AIC                          11937.611
Time:                        14:20:02   BIC                          11973.900
Sample:                             0   HQIC                         11950.637
                               - 3129
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.0269      0.010     -2.724      0.006      -0.046      -0.008
ar.L2         -0.0220      0.010     -2.152      0.031      -0.042      -0.002
ar.L3         -0.0143      0.011     -1.295      0.195      -0.036       0.007
ar.L4         -0.0041      0.010     -0.414      0.679      -0.023       0.015
ar.L5          0.0233      0.010      2.368      0.018       0.004       0.043
sigma2         2.6502      0.030     89.469      0.000       2.592       2.708
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):              9704.31
Prob(Q):                              0.95   Prob(JB):                         0.00
Heteroskedasticity (H):              46.21   Skew:                             0.01
Prob(H) (two-sided):                  0.00   Kurtosis:                        11.63
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

```
ARIMA Model RMSE: 92.86140471031766
```

```
GARCH model summary for AMZN:

                     Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                  Close   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:                8398.29
Distribution:                  Normal   AIC:                          -16788.6
Method:            Maximum Likelihood   BIC:                          -16764.4
                                        No. Observations:                 3128
Date:                Mon, Jun 10 2024   Df Residuals:                     3127
Time:                        14:20:03   Df Model:                            1
                                 Mean Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
mu         1.4872e-03  6.671e-05     22.293 4.273e-110 [1.356e-03,1.618e-03]
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      6.3523e-06  2.654e-11  2.394e+05      0.000 [6.352e-06,6.352e-06]
alpha[1]       0.0500  1.072e-02      4.662  3.131e-06 [2.898e-02,7.102e-02]
beta[1]        0.9300  9.604e-03     96.831      0.000     [  0.911,  0.949]
============================================================================

Covariance estimator: robust
```

```
GARCH Model RMSE: 0.024568778889277627
```

Summary ARIMA Model: The ARIMA model helps in forecasting the stock's closing prices by capturing its trends and patterns.
GARCH Model: The GARCH model helps in modeling and forecasting the volatility (variance) of the stock's returns, providing insights into periods of high and low volatility.\

These models combined can provide a comprehensive analysis of both the expected price levels and volatility of Amazon's stock, making them useful for various forecasting and risk management applications.

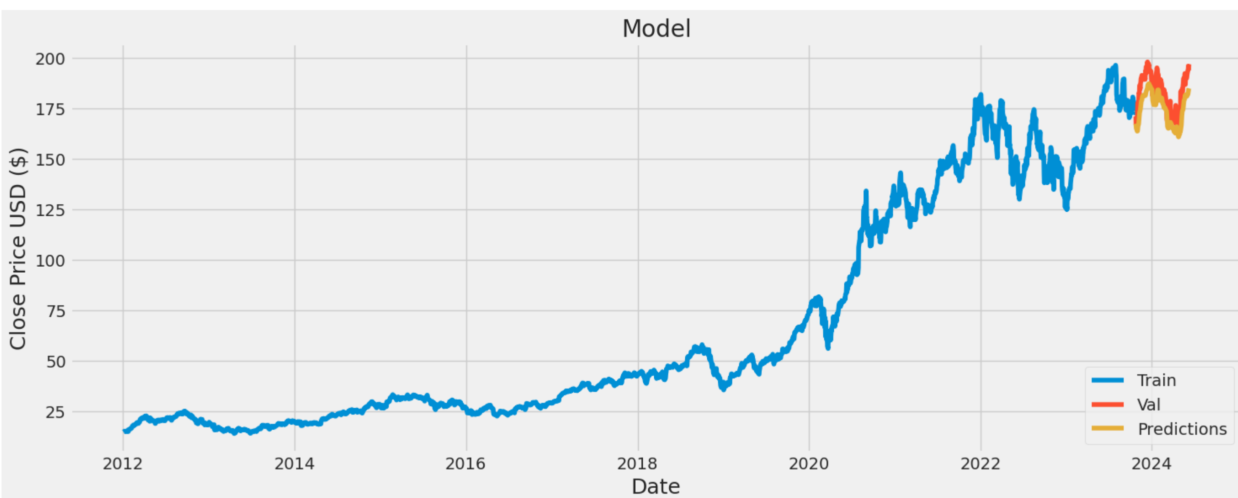
LSTM

```
model.fit(x_train, y_train, batch_size=1, epochs=5)

Epoch 1/5
2913/2913 [==============================] - 176s 59ms/step - loss: 0.0011
Epoch 2/5
2913/2913 [==============================] - 160s 55ms/step - loss: 5.7675e-04
Epoch 3/5
2913/2913 [==============================] - 160s 55ms/step - loss: 3.6685e-04
Epoch 4/5
2913/2913 [==============================] - 161s 55ms/step - loss: 3.0291e-04
Epoch 5/5
2913/2913 [==============================] - 160s 55ms/step - loss: 2.7954e-04
<keras.src.callbacks.History at 0x7fb5fd7c7ac0>
```

```
rmse

5/5 [==============================] - 2s 61ms/step
8.91998581197704
```



	Close	Predictions
Date		
2023-10-25	171.100006	167.478546
2023-10-26	166.889999	166.707062
2023-10-27	168.220001	165.226685
2023-10-30	170.289993	164.125595
2023-10-31	170.770004	163.768784
...	...	...
2024-06-03	194.029999	182.055756
2024-06-04	194.350006	182.875229
2024-06-05	195.869995	183.655685
2024-06-06	194.479996	184.553055
2024-06-07	196.889999	184.996796

156 rows × 2 columns