Financial Data Storytelling

Exploratory data analysis (EDA) is an important part of every data scientist's workflow. EDA allows data scientists to summarize the most important characteristics of the data they're working with. In the case of financial data analysis, this includes generating simple summary statistics such as standard deviation in returns and average returns, visualizing relationships between stocks through correlation heatmaps, generating stock price time series plots, boxplots, and more.

Let's look at three stocks: Amazon (AMZN), Google (GOOGL), and Apple (AAPL) (AAPL). We'll look at how to do some basic exploratory data analysis on these equities, such as creating summary statistics and visualisations, risk and return analysis, and building lagging indicators to better understand stock price movements. This should form a solid foundation for the beginner who wants to get started learning how to analyse financial data in Python. Before we get started, here are some of the tools we'll utilise.

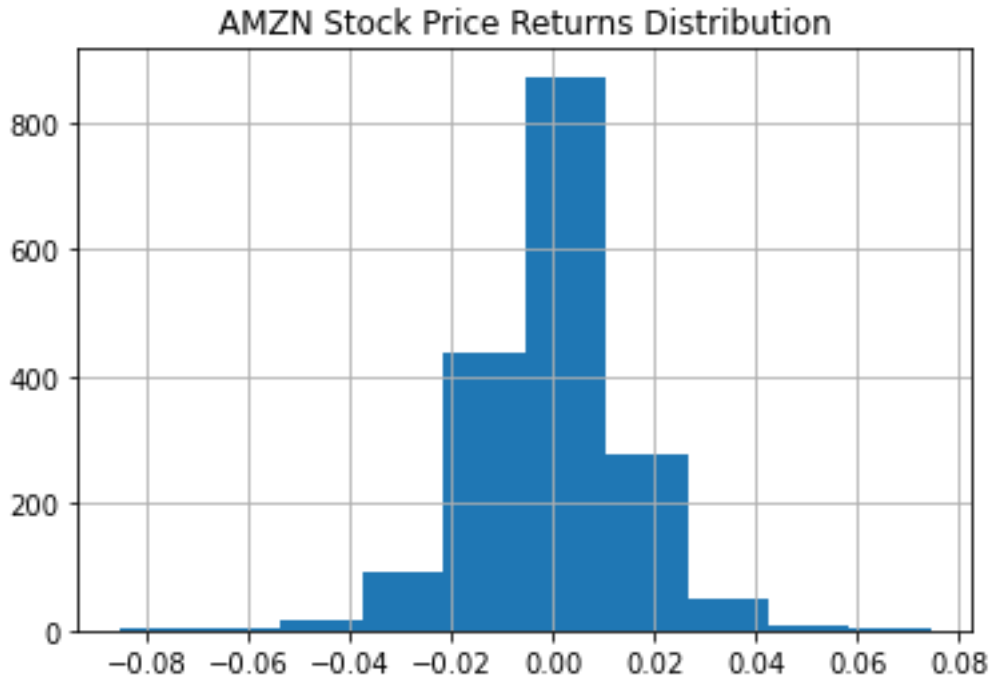
Understanding the uncertainty in stock price volatility necessitates stock risk analysis. This can assist investors in determining which equities to invest in based on their risk tolerance. By characterising the directional trend in stock price movement, we may use moving average calculations to help us make better investment selections.

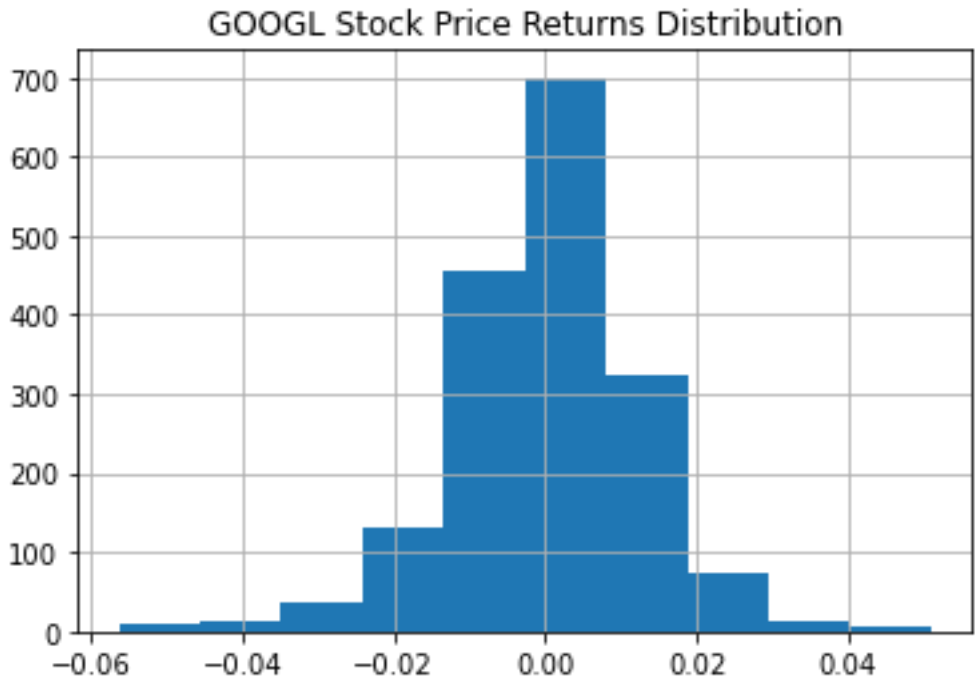
The data frame has columns High, Low, Open, Closed, Volume, and Adjusted Close, as can be seen. These figures are based on stock prices between 9:30 a.m. and 4:00 p.m. during a trading session. Consider the following definitions for each of these columns:

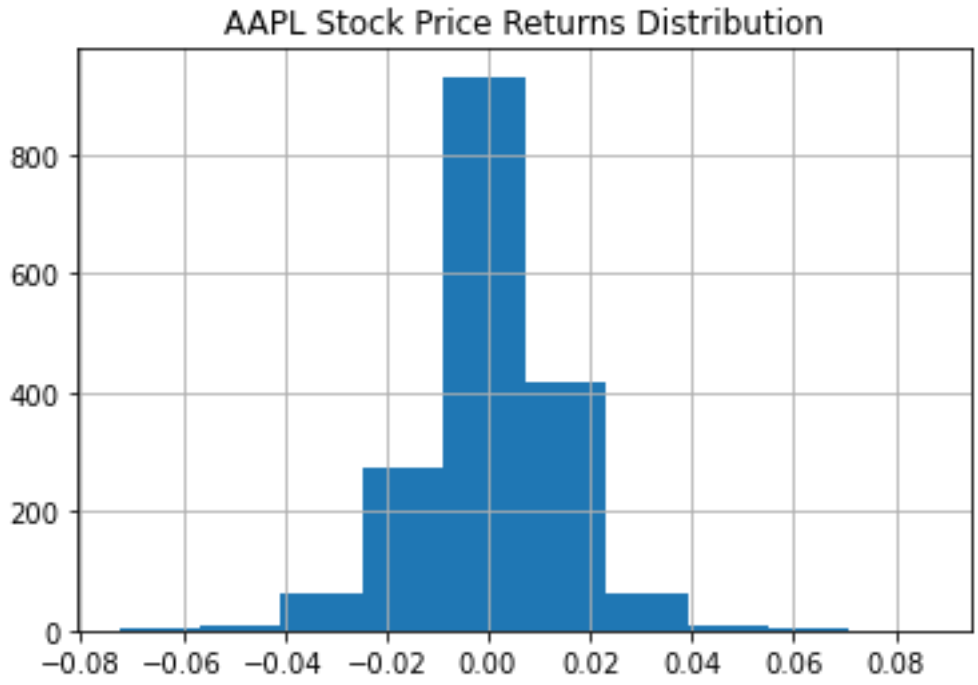
- High price: the highest price of a stock during a trading session
- Low price: a stock's lowest price during a trading session.
- The price of a stock at the end of a trading session is known as the close price.

- The price of a stock at the start of a trading session is known as the open price.
- Adjusted Close: the closing price after stock splits and dividends have been taken into account.

Tech Giant's Returns analysis :

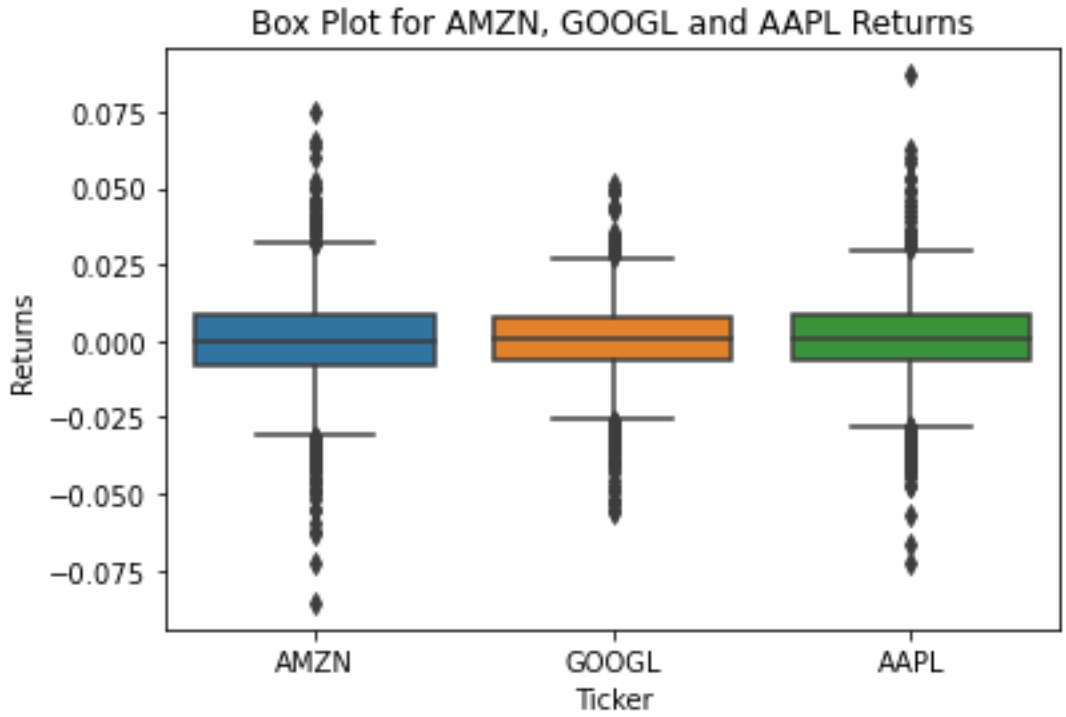


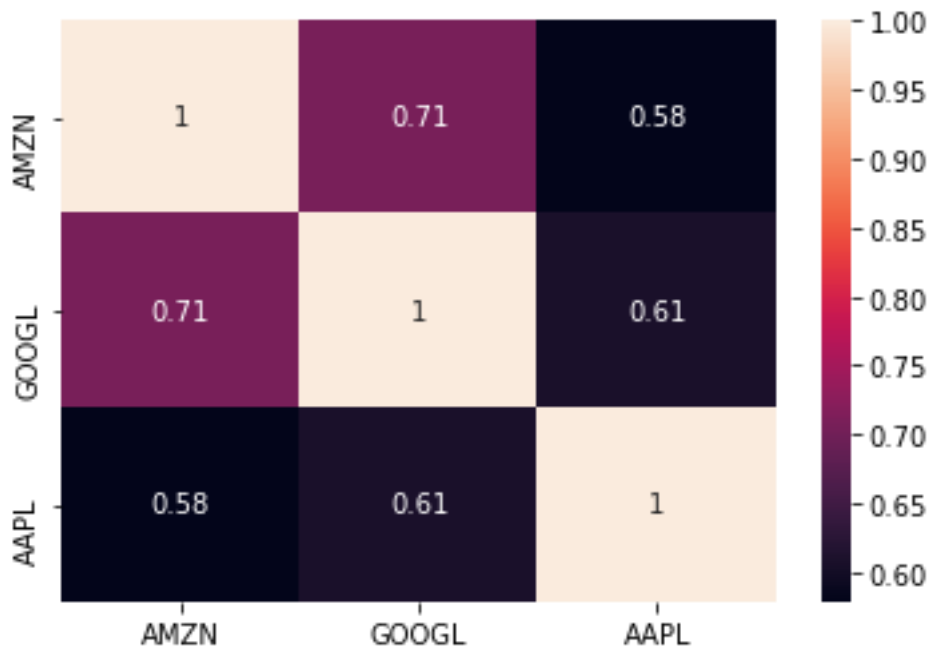




We can see that tech giants have been giving decent results over past few years , but AAPL seems to have a little edge over AMZN in terms of returns. GOOGL seems to chasing AAPL and AMZN at least in the stock returns of past 7 years.

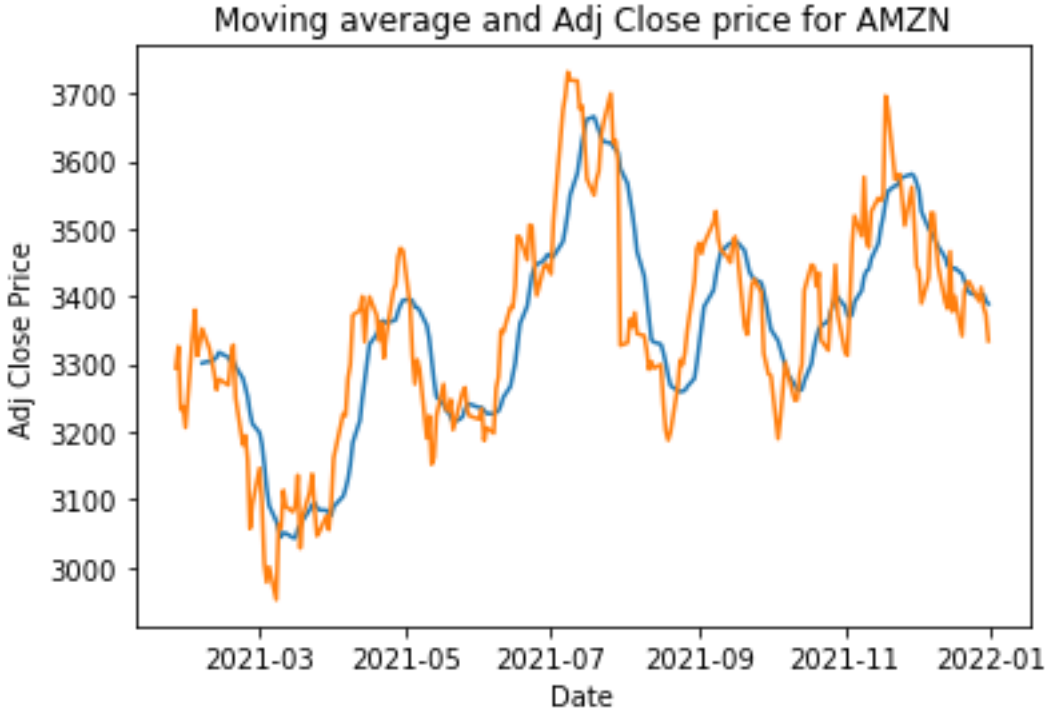
Another useful data visualization is boxplots. Similar to histograms, this is another way to visualize mean, dispersion and skewness in data. In the context of our financial data, it can help us compare the mean returns, the dispersion in returns, and the skewness in returns for each stock, which can help inform investment decisions. First let's combine the returns for each stock into a single data frame:

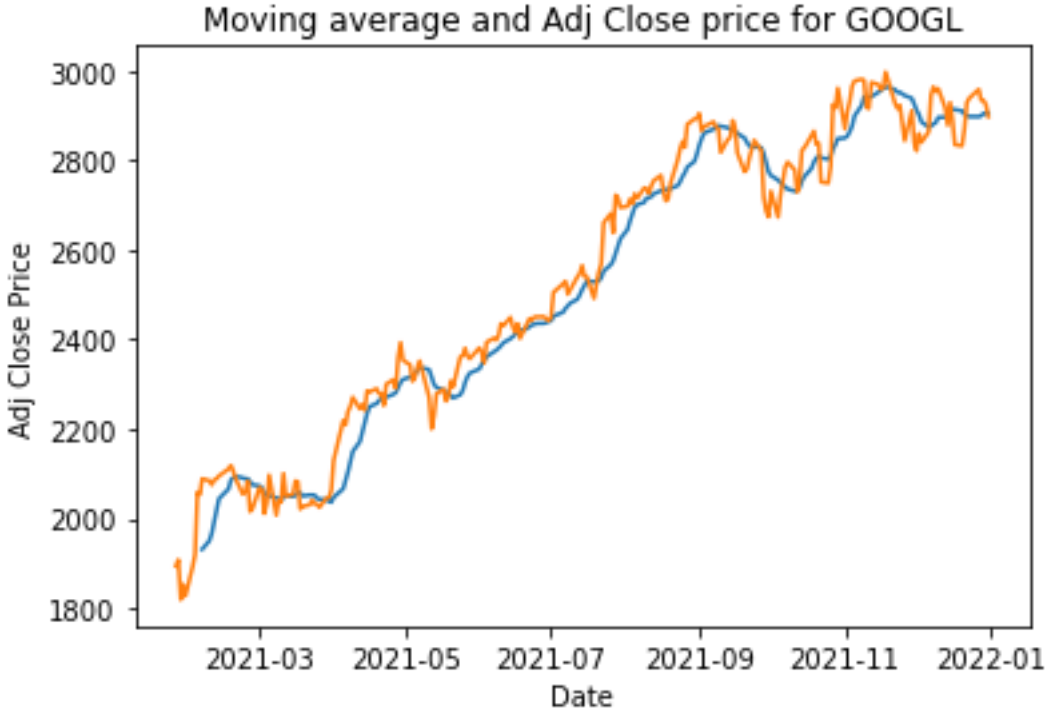


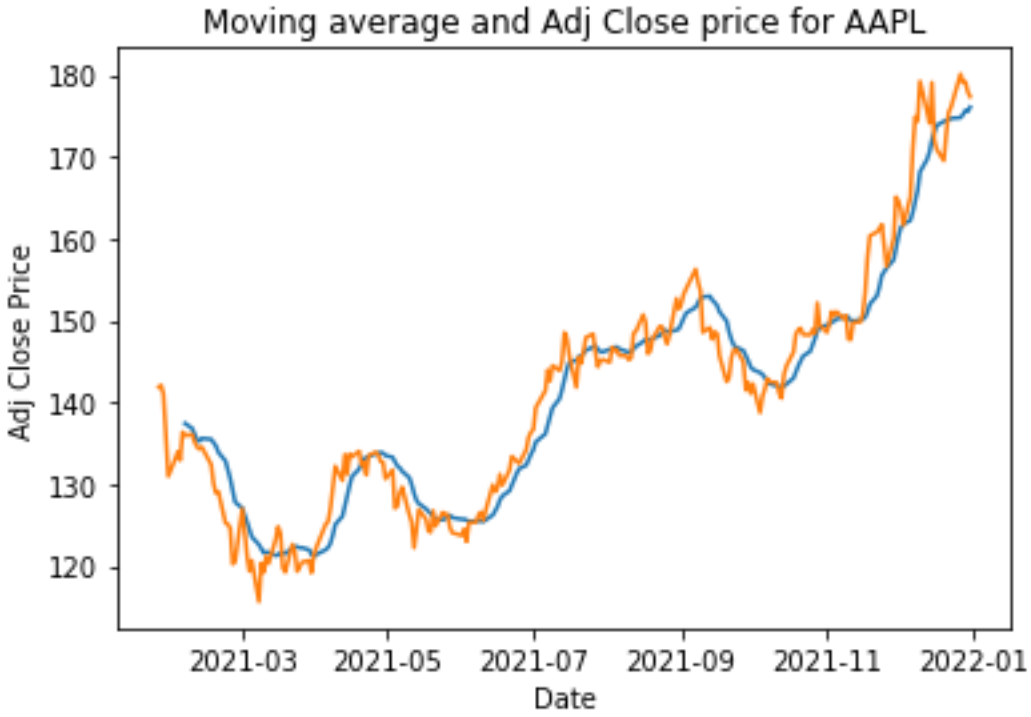


This heatmap shows that each of these stocks have a positive linear relationship. This means that when the daily returns of AMZN increase, AAPL and GOOGL are also likely to increase. The reverse is also true. If AMZN returns decrease, the others are also likely to decrease. A good investment portfolio contains diversified assets. In this context, this means we should select stocks that are not strongly correlated with each other like AAPL, AMZN, and GOOGL. This is because if the returns for one stock dips, your entire portfolio returns will also decrease. In a diversified portfolio with stocks that are uncorrelated, one stock price will not necessarily decrease or increase along with any others.

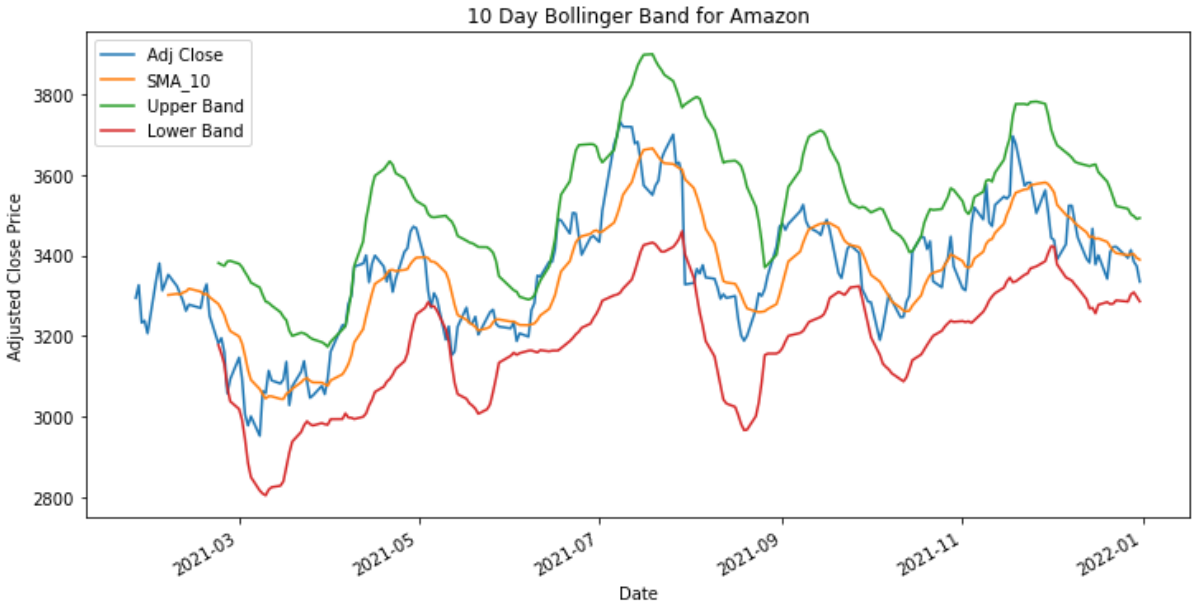
The next calculations we'll walk through are two different kinds of lagging indicators, moving average and Bollinger Band plots. The moving average is a common technique analysts use to smooth out short-term fluctuations in stock prices to understand trends in price direction. Here we'll plot the moving average for AMZN, GOOGL and AAPL.

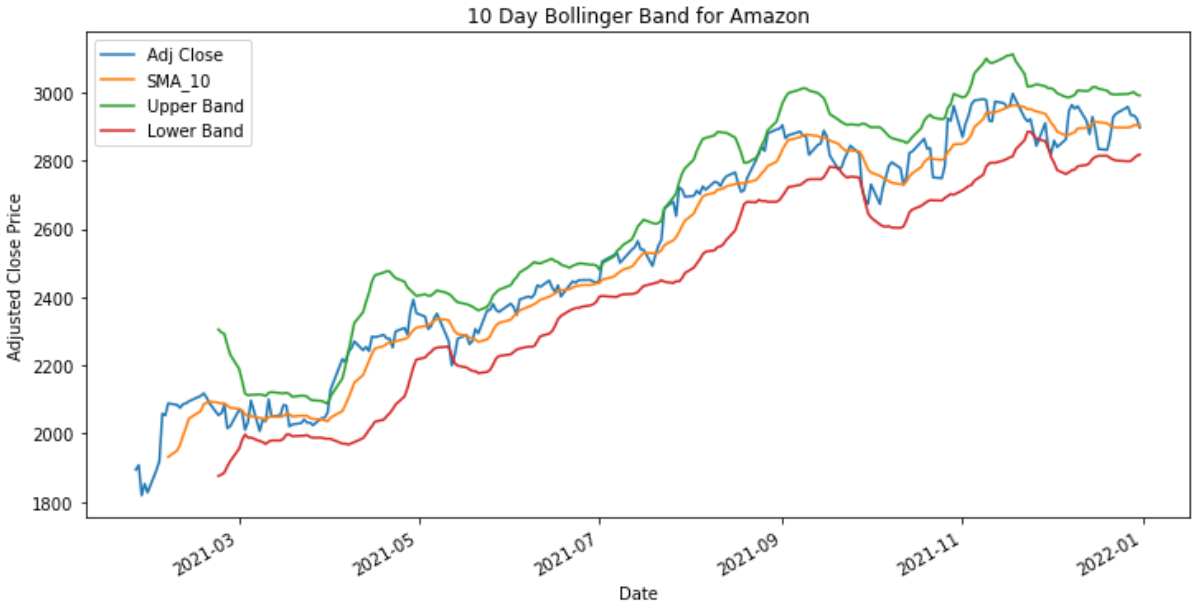

Moving average and Adj Close price for AMZN
Adj Close Price
3700
3600
3500
3400
3300
3200
3100
3000
2021-03
2021-05
2021-07
2021-09
2021-11
2022-01
Date


Moving average and Adj Close price for GOOGL
Adj Close Price
3000
2800
2600
2400
2200
2000
1800
2021-03
2021-05
2021-07
2021-09
2021-11
2022-01
Date



The Bollinger Band plot, which is a means to show the dispersion in the moving average, is the last sort of plot I'll cover. Upper and lower bounds that are two standard deviations distant from the simple moving average define the bands. Traders benefit from this since it allows them to profit from price volatility changes.


10 Day Bollinger Band for Amazon
Adj Close
SMA_10
Upper Band
Lower Band
Adjusted Close Price
3800
3600
3400
3200
3000
2800
2021-03
2021-05
2021-07
2021-09
2021-11
2022-01
Date


10 Day Bollinger Band for Amazon
Adj Close
SMA_10
Upper Band
Lower Band
Adjusted Close Price
3000
2800
2600
2400
2200
2000
1800
2021-03
2021-05
2021-07
2021-09
2021-11
2022-01
Date