Aim:

The aim of this project is to analyze S&P 500 data from the past 93 years (December 1927 to November 2020) through analysis of the daily returns. Effectiveness of technical indicators like the Simple Moving Average (SMA) and Exponential Moving Average (EMA) will be determined based on a 20-day period to generate buy and sell signals.. Based on the analysis I will attempt to present the case for which indicator, if either, is a better predictor based on the long- and short-term scale of investment and how investors can aim to use such indicators to determine the outcome of their investments.

Introduction:

Moving averages are a widely used tool for analyzing asset prices, offering insights into potential future price movements by identifying levels of support and resistance. For instance, when an asset's price crosses below its moving average during a downward trend, it often signals further decline, whereas crossing above it during an upward trend suggests continued growth.

This analysis focuses on using Simple Moving Averages (SMA) and Exponential Moving Averages (EMA) to predict buy and sell signals for the S&P 500 index. Additionally, it aims to provide an overall understanding of the S&P 500 as an investment option by exploring trends and evaluating its performance through the following goals:

1. Calculate daily returns and moving averages.
2. Identify buy and sell signals.
3. Visualize trends in the data to aid decision-making.

Data Description

Data Set:

This analysis is based on historical data of S&P 500 obtained from Nasdaq (https://www.nasdaq.com/market-activity/index/spx/historical). Nasdaq's historical S&P 500 data is widely recognized for its accuracy and comprehensiveness, making it an ideal source for financial trend analysis. It consists of 23,323 data points in a CSV format.Each entry represents important metrics of the S&P 500 during the day, including opening price, closing price, high, low, volume, and adjusted close. The time period covered in this data ranges fromDecember 30, 1927, to November 4, 2020 excluding weekends and other holidays during which the market was closed.

Data Cleaning:

The data frame was analyzed visually for any obvious gaps, missing values or inconsistencies in data type. For more thorough checking the "isnull()" function along with "any()" of the pandas library was utilized to check for missing values within the "Opening Price", "Closing Price" and "Date". It is important to do so, as missing values can cause analysis functions to throw errors and exceptions. Through these checks, we can ensure that the data is complete, and if not, take steps to address missing values. In our case the “isnull()” and “any()” functions returned “False” for the given fields indicating that the data frame had no missing values or values with inconsistent data type. Such checks were not performed on “High,” “Low,” and “Volume” data. These metrics were excluded as this analysis focuses on trends over daily closing prices, which are more reflective of long-term market movement. Volume was omitted because this study is focused on price trends rather than trading activity.

Approach and Methodology:

The analysis began by ensuring the dataset was clean and consistent. The first step involved plotting the **Closing Price** against dates to observe the overall trends in the S&P 500’s historical performance. While the closing price provides a general sense of market direction over time, it does not account for intraday fluctuations that short-term investors might prioritize. To address this, **daily returns** were calculated to better understand day-to-day price changes, offering a clearer view of short-term market behavior. The following formula was used:

$$\text{Daily Return} = \frac{\text{Current Day's Close} - \text{Previous Day's Close}}{\text{Previous Day's Close}} \times 100$$

1. **Current Day's Close**: The closing price of the asset for the current day.
2. **Previous Day's Close**: The closing price of the asset for the previous day.
3. The result is multiplied by 100 to express the return as a percentage.

Example:

If the closing price on Day 1 is $100, and on Day 2 it is $105:

$$\text{Daily Return} = \frac{105 - 100}{100} \times 100 = 5\%$$

This shows a 5% gain in the asset's price from Day 1 to Day 2.

The next step was calculating the **Simple Moving Average (SMA)**, a technical indicator used to identify support or resistance levels based on historical prices. The SMA is determined by averaging the closing prices over a specified time period, providing a lagging indicator of price trends. For this analysis, a 20-day SMA was chosen, balancing the need to capture short-term trends with computational efficiency, as analyzing the entire dataset was resource-intensive. Shorter SMAs like the 20-day average are particularly effective in detecting recent price movements and trends.

Formula of Simple Moving Average

$$\text{Simple Moving Average} = \frac{(A_1 + A_2 + \ldots\ldots + A_n)}{n}$$

$(A_1, A_2 \ldots\ldots A_n)$ = Prices

n = The number of total periods

The analysis then incorporated the **Exponential Moving Average (EMA)** to complement the SMA. Unlike the SMA, the EMA assigns more weight to recent price data, making it more responsive to current market conditions. This attribute allows the EMA to reflect recent events more effectively, offering a sharper and more immediate perspective on price trends.

Exponential moving average is calculated by the following formula:

$$\text{EMA} = (P * \alpha) + (\text{Previous EMA} * (1 - \alpha))$$

P = Current Price

α = Smoothing Factor = $\frac{2}{1+N}$

N = Number of Time Periods

Finally, **buy** and **sell signals** were generated based on both SMA and EMA. A **buy signal** was identified when the **Closing Price** exceeded the moving average, while a **sell signal** occurred when the **Closing Price** fell below it. Additionally, the interaction between the SMA and EMA provided deeper insights: when the EMA crossed above the SMA, it indicated a potential uptrend, whereas an EMA crossing below the SMA suggested a possible downtrend. By combining these indicators, the analysis provided a robust framework for interpreting market trends and making informed investment decisions.

Results:

From S&P 500 plot:

The S&P 500 historical chart illustrates a significant long-term upward trend, showcasing the growth and resilience of the stock market over the past century. While the overall trajectory is positive, the chart reveals periods of volatility and stagnation associated with major economic events. Notable declines can be observed during the Great Depression in the 1930s, the Dot-com Bubble in the early 2000s, and the Global Financial Crisis of 2008–2009, reflecting the impact of these crises on market performance. However, from the 1980s onward, the market experienced accelerated growth, driven by technological advancements and globalization. This trend steepens notably after the 2008 financial crisis, as the market recovered strongly due to robust monetary policies.

The resilience of the S&P 500 is evident in its ability to recover from sharp declines and continue its upward trajectory, rewarding long-term investors who withstand short-term fluctuations. The chart also highlights recent peaks in market performance prior to 2020, which could be contextualized by events such as the COVID-19 pandemic. Despite the short-term volatility that likely characterizes daily returns during these turbulent periods, the long-term growth emphasizes the benefits of patience and a forward-looking investment strategy.

From Daily Return plot:

Daily returns were plotted for the past year (November 5, 2019 – November 4, 2020) using the formula stated above. A one-year interval was chosen to ensure clarity in the graph and avoid overcrowding. Overall, the graph demonstrates that the S&P 500 exhibited relatively consistent behavior with limited volatility, except for the period from late February to mid-April. During this time, wide fluctuations in returns—both positive and negative—can be observed, presumably due to the uncertainty that gripped financial markets and the broader economy at the onset of the novel coronavirus pandemic.

This observation aligns with the S&P 500 price graph for the same period, which shows a sharp decline of approximately 1,200 points by mid-March, followed by a recovery of the losses during the second quarter of 2020.

SMA and EMA Analysis-Based Buy and Sell Signals:

A 20-period SMA and EMA were computed and plotted alongside the daily price data to illustrate how these indicators align with price trends and to evaluate their effectiveness in generating buy and sell signals. In the analysis, green markers were used to indicate “buy” signals, while red markers represented “sell” signals. It was observed that neither SMA nor EMA consistently provided accurate predictions for buy or sell signals. Of the 20 signals generated, approximately half were inaccurate, with some buy signals occurring just before a decline in price, and sell signals appearing just before an upward trend. This inconsistency is expected, as both SMA and EMA are lagging indicators that reflect historical price trends rather than predicting future price movements. The results highlight that SMA and EMA are not reliable when used in isolation as standalone tools for investment decisions. To enhance their predictive accuracy, it is recommended to combine them with other technical indicators, such as the Relative Strength Index (RSI), to gain a more comprehensive understanding of market conditions and the viability of trades.

Conclusions:

Based on our analysis, SMA stands as a better predictor of price trends for long-term investors looking to acquire shares or stocks of a company for a year or more. Since SMA weighs all data points equally for a given time period it is more likely to not be impacted by the day to day noise that is part of trading and give a more accurate picture of pricing trends over the course of several months to years. EMA, on the other hand, being a more sensitive technical indicator to the latest data, is a worthwhile analysis to be performed for short-term day-to-day traders. Ideally, our recommendation is that investors (especially short term investors) use EMA in tandem with the SMA and pay close attention to points where the two averages intersect. In a heavily bullish market EMA intersecting SMA during a downward trend for both would signal losses ahead and vice versa.

As responsive as EMA is, it doesn't always capture all spikes and dips in the stock price. Moving averages inherently smooth out the pricing fluctuations so there certainly can be some changes in prices that aren't strong enough to have an impact on the EMA, especially EMAs that are performed over longer time periods. Furthermore, moving averages are a lagging indicator; any insights drawn from them are based on the assumption that the past (and present) trends will continue. Unexpected events that cause stock prices to rise or decline quickly, and the intensity of such events, may not be accurately captured by EMA. Investors should combine technical indicators like EMA with fundamental analysis, industry trends, and market news to make more informed decisions.

Indescribable events, while rare, do occur from time to time and can potentially have a major and lasting impact on the price of a stock and hence its attractiveness as an investment. However, the effects of such events and their extent aren't something that EMA can predict. A recent example of such an event is the effect the deepseek AI model had on NVDA trading stock prices at the end of Jan 2025. Deepseek AI was developed in a Chinese lab and proved to be a better, cheaper and more efficient version of OpenAI's ChatGPT despite lack of access to the most recent GPUs that Nvidia was providing for AI development within the US. Deep seek performed comparable or better than any other AI model available. NVDA stock lost about 24% of its value in one day on Jan 27th 2025. However a 20 period EMA was relatively flat going into Jan 28th signaling some stability and potentially a point for short term investors to get in the action. By the close of Jan 28th NVDA was up 8% in just one day. Since the EMA isn't as responsive to such events as the price itself therefore it can give short term investors an idea of when it is an appropriate time to acquire an instrument like Nvidia.

As far as the future price predictions are concerned moving averages provide probabilistic insights, they do not guarantee future price movements. However, since EMA and SMA for the S&P 500 index have continued to increase in value and recover losses they incurred it can be predicted that the price will continue to rise over a longer time period. In the data used for this analysis the last point is on Nov 04th 2020. In this case the EMA had just crossed below SMA around 10/23/2020. However the SMA was flattening up and EMA seemed to be on an upward trajectory. If this current trend continues and EMA crosses above SMA again, a short-term rally could follow over the next 30 days. This also is the time going into the last trading month of the year during which the stock value typically rises so the projections of the study are congruent with the seasonality.

While this study primarily focused on SMA and EMA, it's important to recognize that other technical indicators—such as the Relative Strength Index (RSI), Moving Average Convergence

Divergence (MACD), and Bollinger Bands—can also contribute valuable insight into stock price behavior. These tools can help identify momentum, overbought/oversold conditions, and volatility, which moving averages alone may not fully capture. Integrating these alongside EMA could improve the accuracy of buy/sell signals, especially for short-term strategies.

Furthermore, analyzing the **correlation between these additional indicators and stock prices**—as well as their interaction with EMA/SMA—could enhance forecasting reliability. For example, a rising EMA that coincides with an RSI below 70 (not overbought) may strengthen confidence in a bullish trend.

Future research could explore these correlations quantitatively, using statistical methods such as Pearson correlation coefficients or regression models, to determine how strongly these variables align with price movements and with each other. This multi-indicator approach may yield more robust trading signals and deeper insights into market dynamics.

Implementation:

Dependencies:

The following tools and libraries were used to implement the analysis:

- **Microsoft Excel**: For initial data review and organization.
- **Python 3.9**: The primary programming language for processing and analysis.
- **Matplotlib 3.9.2**: Used to create visualizations and plot trends in the data.
- **Pandas 2.2.2**: Utilized for efficient data manipulation, cleaning, and computation.