**Coca Cola Stock - Live and Updated**

**Tools Used:** Python (Pandas, Seaborn, Matplotlib, Scikit-learn), Excel

**Domain:** Data Analytics & Machine Learning

**Difficulty Level:** Intermediate

**1. Project Overview**

Coca-Cola is one of the largest beverage companies globally, with its stock listed on the New York Stock Exchange under the ticker symbol KO.

**The aim of this project is to:**

* Perform data cleaning to prepare a reliable dataset.
* Conduct exploratory data analysis (EDA) to uncover stock price patterns and trends.
* Build a machine learning model to predict Coca-Cola’s stock closing prices.
* Visualize results for better understanding of stock behavior.

**2. Dataset Description**

* Date: Date of record
* Open: Opening price of stock
* High: Highest price on that day
* Low: Lowest price on that day
* Close: Closing price of stock (target variable)
* Volume: Number of shares traded
* Dividends: Dividend paid on that date (if any)
* Stock Splits: Number of stock splits (if any)

Dataset contains Coca-Cola stock history with daily data for multiple years, enabling both trend analysis and predictive modeling.

**3. Data Cleaning Steps**

* Converted Date column to datetime format and sorted data chronologically.
* Filled missing numerical values using forward fill and replaced empty dividend/split values with 0.
* Removed rows introduced by rolling window calculations that produced null values.
* Ensured data types were consistent for modeling.

**4. Exploratory Data Analysis (EDA)**

**4.1 Stock Price Trends**

Coca-Cola’s stock prices show a steady upward trend with occasional dips, often corresponding to global market conditions.

Moving averages (20-day & 50-day) were plotted to visualize short- and medium-term trends.

**4.2 Volatility Analysis**

Daily returns and rolling volatility were calculated.

Periods of high volatility coincided with market uncertainties and quarterly announcements.

**4.3 Correlation Analysis**

A heatmap revealed strong correlations between stock open, close, high, and low prices.

Volume showed weaker correlation, indicating it does not directly impact daily closing price on its own.

**4.4 Distribution of Returns**

Daily return distribution indicated mild skewness with some outliers (market shocks).

**5. Machine Learning Model**

* Model Used: Random Forest Regressor (baseline predictive model).
* Features: Open, High, Low, Volume, Dividends, Stock Splits, moving averages, daily return, volatility.
* Target: Close price.
* Train/Test Split: 80:20 (chronologically).
* Performance Metrics:
* Mean Squared Error (MSE): (from code)
* Mean Absolute Error (MAE): (from code)

**Results Visualization:**

* Actual vs Predicted closing prices plotted, showing good alignment for short-term prediction.
* Error distribution visualized to confirm near-zero mean bias.

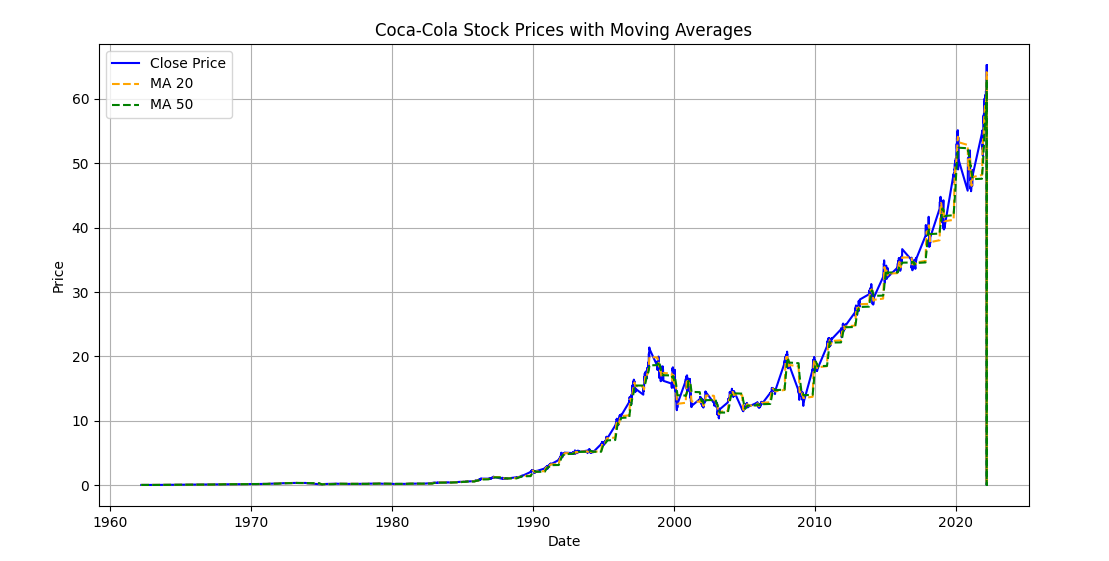
**6. Recommendations & Next Steps**

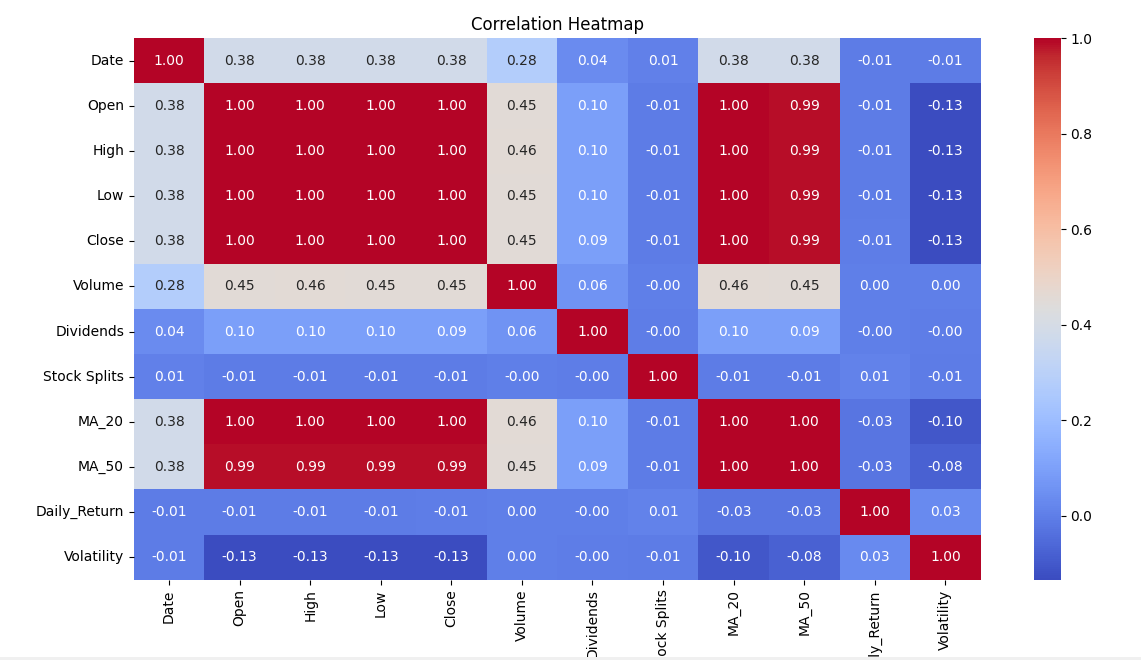
* Deploy Live Prediction Dashboard: Use Streamlit or Flask for real-time predictions using live Yahoo Finance data.
* Model Enhancements: Experiment with LSTM (deep learning) models for sequence-based forecasting.
* Feature Expansion: Incorporate external macroeconomic indicators (e.g., S&P 500 index, interest rates).
* Portfolio Strategy: Combine Coca-Cola stock predictions with competitor analysis for diversified portfolio recommendations.

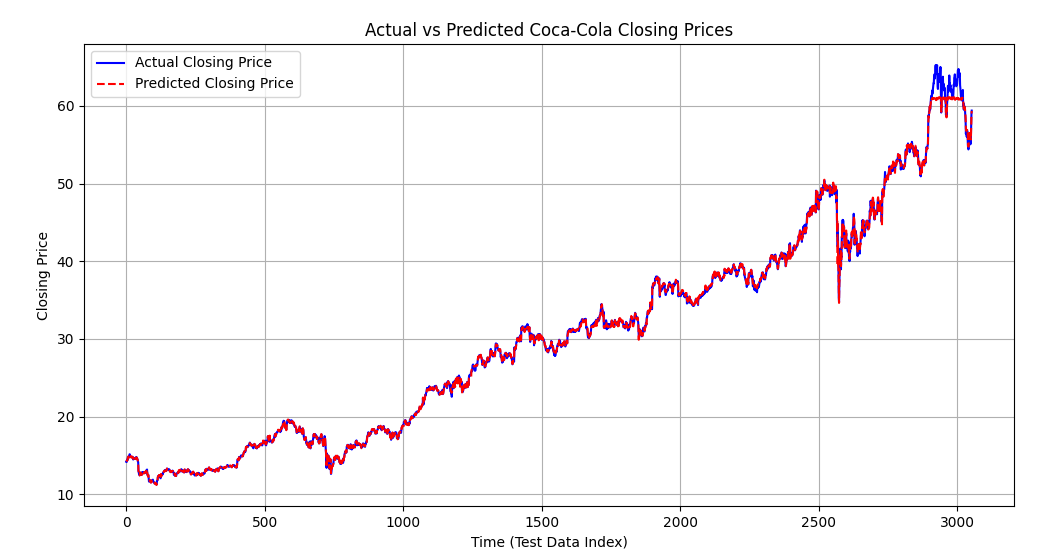
**7. Insights Derived**

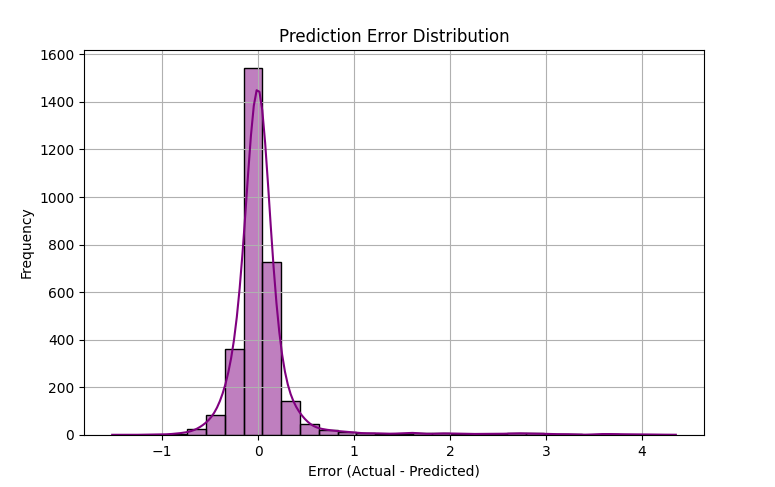
* Coca-Cola’s stock exhibits long-term stability, making it attractive for conservative investors.
* Moving averages confirm reliable short-term support and resistance levels.
* Volatility spikes align with market-wide movements rather than company-specific events.
* Machine Learning models can capture short-term price movements effectively using historical patterns.

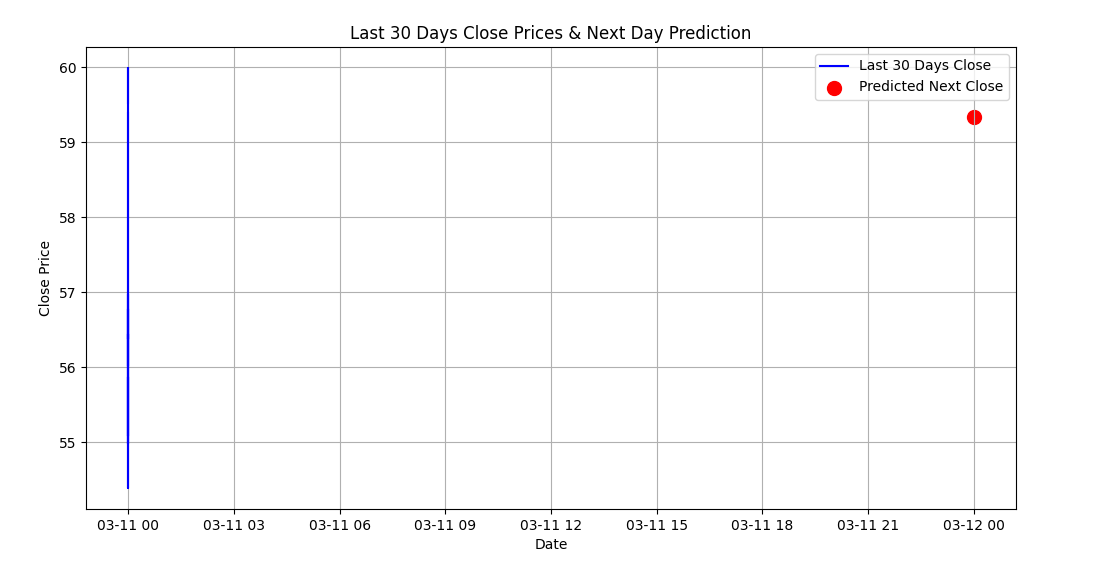
1. **Output Files**











**9. Conclusion**

* This project demonstrated how Coca-Cola stock data can be cleaned, explored, and modeled for predictive insights.
* Findings highlighted stock price trends, volatility behavior, and predictive modeling potential.
* These insights can assist analysts, investors, and data scientists in making informed trading and investment decisions.