**Project Proposal: Coca-Cola Stock Analysis**

**Project Title:** Coca-Cola Stock Analysis and Prediction

**Tools & Technologies:**

- Python

- Machine Learning (ML)

- SQL

- Excel

- Jupyter Notebook

- Pandas

- Matplotlib & Seaborn

- Scikit-learn

**Domain:**

Data Analysis & Machine Learning

**Dataset:**

The dataset can be downloaded from the provided link. The dataset includes historical stock prices and related information for Coca-Cola.

**Objective:**

- Predict Coca-Cola's stock prices (e.g., closing price) and analyze trends using historical data.

- Develop a machine learning model for accurate stock price prediction.

- Create visualizations and perform Exploratory Data Analysis (EDA) to derive insights.

**Methodology:**

1. Problem Definition:

- Define the project's objectives and scope.

- Identify the key metrics and deliverables.

2. Data Collection:

- Use the Yahoo Finance API to fetch historical stock data from 2015 to the present.

- Load the data into a Pandas DataFrame for further analysis.

3. Data Cleaning:

- Handle missing values and outliers.

- Perform data imputation or deletion as necessary.

4. Feature Engineering:

- Create new features such as moving averages, daily returns, and volatility.

- Ensure features are scaled and normalized.

5. Exploratory Data Analysis (EDA):

- Use descriptive statistics to summarize the data.

- Visualize trends and relationships using line plots, histograms, and heatmaps.

6. Data Splitting:

- Split the data into training and testing sets for model evaluation.

7. Model Training:

- Train a Random Forest model (or another suitable ML model) for initial predictions.

- Evaluate model performance using metrics such as Mean Absolute Error (MAE) and Mean Squared Error (MSE).

**Conclusion/Deliverables:**

1. **Stock Performance**: Coca-Cola's stock has shown consistent growth over the years, with significant increases in stock price, dividends, and market capitalization. The stock has been a reliable investment option for long-term investors.

2. **Trend Analysis**: The EDA revealed clear trends in the stock's closing prices, moving averages, and volatility. The moving averages, such as the 20-day and 50-day moving averages, helped identify short-term and long-term trends in the stock prices.

3. **Predictive Model**: The machine learning model, specifically the Random Forest Regressor, provided accurate predictions for the stock's closing prices. The model's performance was evaluated using metrics like Mean Absolute Error (MAE) and Mean Squared Error (MSE), and it demonstrated good predictive capabilities.

4. **Feature Importance**: Feature engineering played a crucial role in enhancing the model's performance. Features like moving averages, daily returns, and volatility were significant in predicting the stock prices.