**Data-Driven Stock Analysis: Organizing, Cleaning, and Visualizing Market Trends**

**1. Introduction**

In the ever-evolving stock market, data-driven insights are crucial for investors and analysts. This project aims to analyze the performance of Nifty 50 stocks over the past year by organizing, cleaning, and visualizing market trends. It leverages Python, Pandas, SQL, Power BI, and Streamlit to provide interactive dashboards for informed decision-making.

**2. Problem Statement**

The Stock Performance Dashboard provides a comprehensive visualization of daily stock data, including open, close, high, low, and volume values. It cleans and processes the data, generates key performance insights, and visualizes trends for better investment decisions.

**3. Business Use Cases**

* **Stock Performance Ranking:** Identifies the top 10 best- and worst-performing stocks.
* **Market Overview:** Summarizes average stock performance and categorizes green vs. red stocks.
* **Investment Insights:** Highlights stocks with consistent growth or decline.
* **Decision Support:** Provides insights into average prices, volatility, and stock behavior.

**4. Methodology**

**4.1 Data Extraction and Transformation**

* Data is provided in YAML format, organized monthly.
* Extracts data into CSV files categorized by stock symbols.
* Produces 50 structured CSV files for further analysis.

**4.2 Data Analysis and Visualization**

**Python Data Frame for Key Metrics:**

* Identifies the top 10 best- and worst-performing stocks based on yearly return.
* Computes the number of green vs. red stocks, average price, and average volume.

**Volatility Analysis:**

* Calculates daily returns and their standard deviation.
* Bar chart visualizing the top 10 most volatile stocks.

**Cumulative Return Over Time:**

* Computes cumulative return from the start to the end of the year.
* Line chart displaying top 5 performing stocks over time.

**Sector-wise Performance:**

* Groups stocks by sector and calculates average yearly return.
* Bar chart showing sector performance.

**Stock Price Correlation:**

* Computes correlation coefficients between stock prices.
* Heatmap visualizing stock price correlations.

**Top 5 Gainers and Losers (Month-wise):**

* Identifies the top 5 gainers and losers for each month.
* Monthly bar charts showcasing performance trends.

**5. Implementation**

**5.1 Tools and Technologies**

* **Languages:** Python
* **Database:** MySQL/PostgreSQL
* **Visualization Tools:** Streamlit, Power BI
* **Libraries:** Pandas, Matplotlib, SQLAlchemy

**5.2 Project Deliverables**

* **SQL Database:** Contains cleaned and processed data.
* **Python Scripts:** For data extraction, cleaning, and analysis.
* **Power BI Dashboard:** Visualizations for stock performance.
* **Streamlit Application:** Interactive real-time data dashboard.

**6. Results and Insights**

* A functional dashboard displaying the best- and worst-performing stocks.
* Interactive visualizations for clear market trend analysis.
* Insights into volatility, sector performance, and cumulative returns.

**7. Conclusion**

This project effectively organizes, cleans, and visualizes stock market data, helping investors and analysts make informed decisions. With the use of Power BI and Streamlit, users can interactively explore stock trends and gain actionable insights.