Stock Price Time Series Forecasting

Summary

This report details the process and findings of a time series forecasting project focused on predicting the adjusted closing prices of five diverse stocks: MARA, SOXL, TD_BANK, NVDA, and MANULIFE. The project involved data collection and exploration, time series decomposition to understand underlying patterns, and the implementation and comparison of four forecasting models: ARIMA, Exponential Smoothing (ETS), Prophet, and LSTM. Exploratory Data Analysis revealed varying trends and volatility across the stocks, with some requiring differencing for stationarity. Decomposition highlighted trend and residual components more strongly than clear seasonality. Model performance varied significantly by stock, with Prophet often showing competitive results on the test set, while LSTM demonstrated potential but also significant instability, particularly in long-term forecasts for certain stocks. The results underscore the importance of model selection tailored to specific data characteristics and the critical need for hyperparameter tuning.

---

1. Dataset and Exploratory Data Analysis (EDA)

1.1. Dataset

The dataset consists of daily stock price information for five companies obtained from the Stock_Data.xlsx file. The analysis focused on the 'Adj Close' (Adjusted Close) price, treated as the target variable for forecasting, using the 'Date' column as the time index. The data spans roughly from early 2024 to April 2025, resampled to business day frequency ('B'). The stocks selected represent different sectors and volatility profiles:

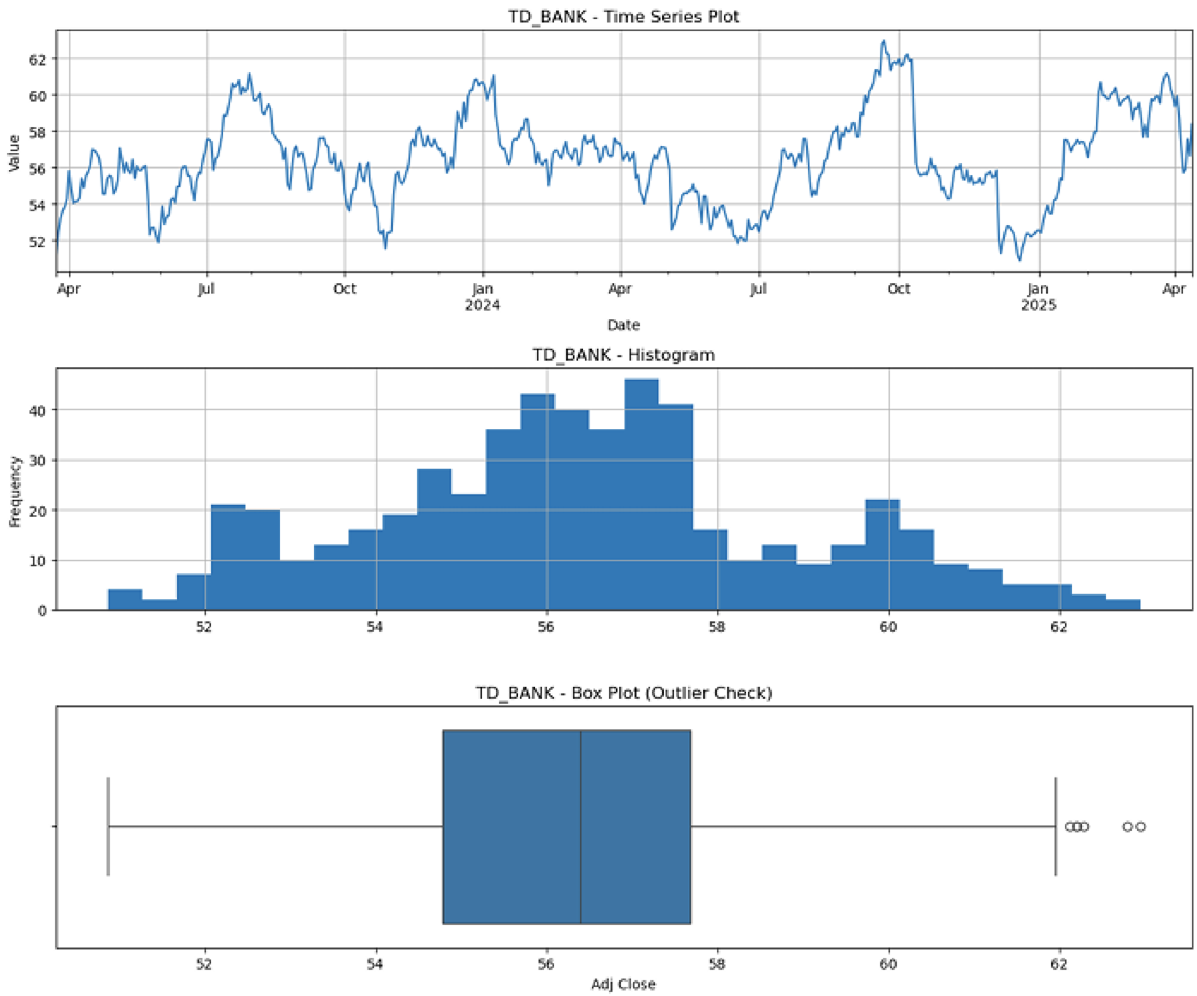
- **MARA (Marathon Digital Holdings Inc.):** Bitcoin Mining (High Volatility)
- **SOXL (Direxion Daily Semiconductor Bull 3X Shares):** Leveraged Semiconductor ETF (High Volatility)
- **TD_BANK (Toronto-Dominion Bank):** Banking/Finance (Moderate Volatility)
- **NVDA (NVIDIA Corporation):** Technology/Semiconductors (High Growth/Volatility)
- **MANULIFE (Manulife Financial Corporation):** Insurance/Finance (Moderate Volatility)

1.2. Exploratory Data Analysis

EDA was performed on each stock's time series data.

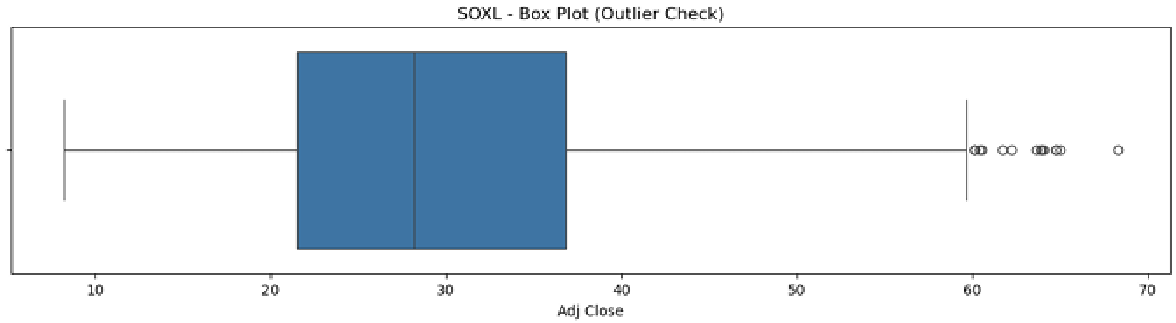
- **Visualization:** Time series plots revealed distinct behaviors. NVDA exhibited a strong upward trend over the period. MARA and SOXL displayed high volatility with significant peaks and troughs. TD_BANK and MANULIFE showed more moderate trends and fluctuations.

Figure 1: Example Time Series Plots for TD_Bank (Trend & Volatility)



- **Missing Values:** Data was resampled to business day frequency. Missing values resulting from non-trading days were handled using forward fill (method='ffill'), assuming the price remained constant from the previous business day. This is a common practice but could potentially smooth over sharp weekend/holiday gaps.
- **Outliers:** Box plots were generated to identify potential outliers (values significantly distant from the interquartile range). Outliers were visually detected in SOXL, TD_BANK, and MANULIFE. For this project, outliers were noted but not explicitly removed or adjusted, which could impact model fitting, particularly for ETS and ARIMA, and scaling for LSTM.

Figure 2: Example Box Plot for SOXL showing outliers



- **Stationarity:** The Augmented Dickey-Fuller (ADF) test was used to check for stationarity.
    - **Stationary (p <= 0.05):** MARA, TD_BANK
    - **Non-Stationary (p > 0.05):** SOXL, NVDA, MANULIFE
    - **Differencing:** For the non-stationary series, the ADF test on the first difference confirmed stationarity (p-value ≈ 0.0000). This indicates that differencing (specifically, d=1) is necessary for models like ARIMA that assume stationarity. ETS models can handle trends directly.

- **External Factors:** Stock prices are inherently influenced by external factors not explicitly included as features in this analysis (except potentially implicitly in Prophet). These include:
    - **Market Events:** Company-specific news (earnings, product launches), sector trends.
    - **Economic Indicators:** Interest rates, inflation, GDP growth.
    - **Holidays:** Reduced trading volume, potential pre/post-holiday effects.
    - **Global Events:** Geopolitical situations, pandemics (though less relevant for this specific timeframe).
    These factors contribute to the 'noise' or residuals in the models and can cause sudden shifts not captured by historical price patterns alone.

2. Time Series Decomposition

Time series decomposition using statsmodels.seasonal_decompose was performed to separate each stock's adjusted close price into three components: Trend, Seasonality, and Residuals. A period of 30 was used, aiming to capture potential monthly cyclical patterns, although true monthly seasonality is often weak in individual stocks compared to broader market indices or retail data.

- **Components:**
    - **Trend:** Captures the long-term direction or level change in the stock price (e.g., the clear upward trend in NVDA, the fluctuating trend in MARA).
    - **Seasonality:** Represents patterns repeating over the specified period (30 days). Visual inspection suggested relatively weak and possibly noisy seasonal components for most stocks at this frequency.
    - **Residuals:** The remaining 'noise' after removing the trend and seasonal components. Represents unpredictable fluctuations or events.

- **Additive vs. Multiplicative Decomposition:**
    - **Additive (Observed = Trend + Seasonal + Residual):** Assumes the magnitude of seasonality and residuals is constant regardless of the price level.
    - **Multiplicative (Observed = Trend * Seasonal * Residual):** Assumes seasonality and residual magnitude scales proportionally with the price level. For stock prices, volatility often increases as the price level rises, making multiplicative decomposition theoretically more appropriate. Visually comparing the plots (e.g., MANULIFE below), the multiplicative model sometimes shows a more stable residual variance relative to the trend.

Figure 3: Additive Decomposition for MANULIFE

```
--- Additive Decomposition for MANULIFE (Period=30) ---
```

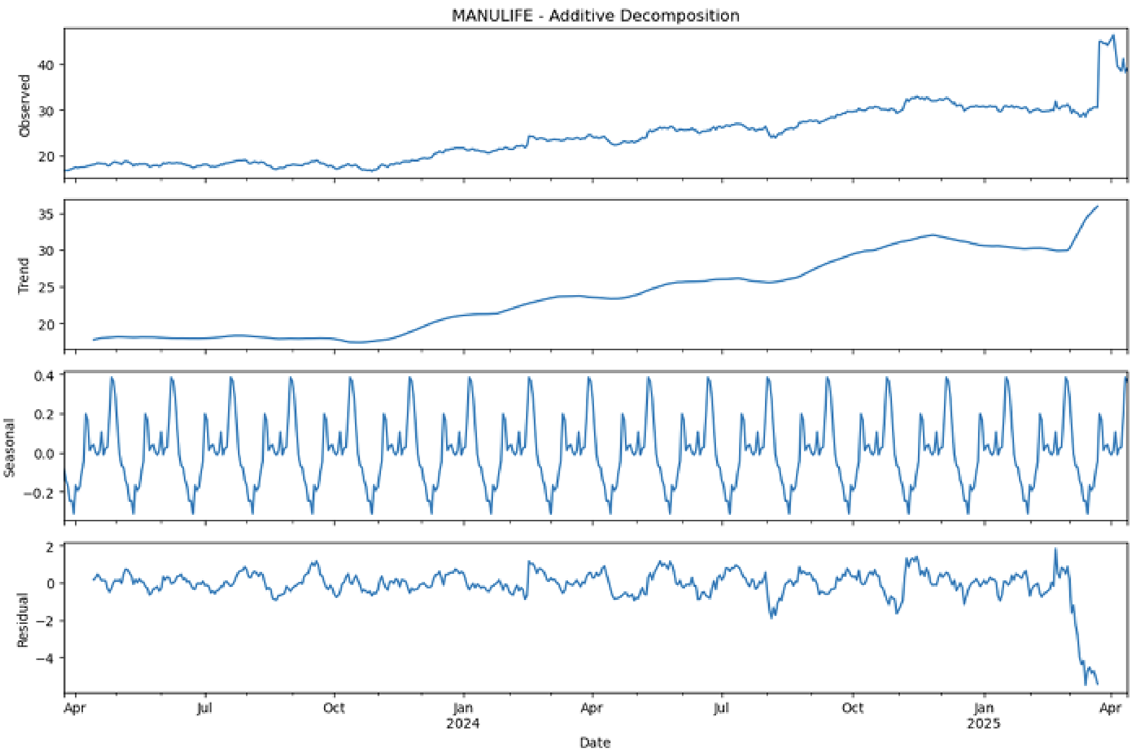
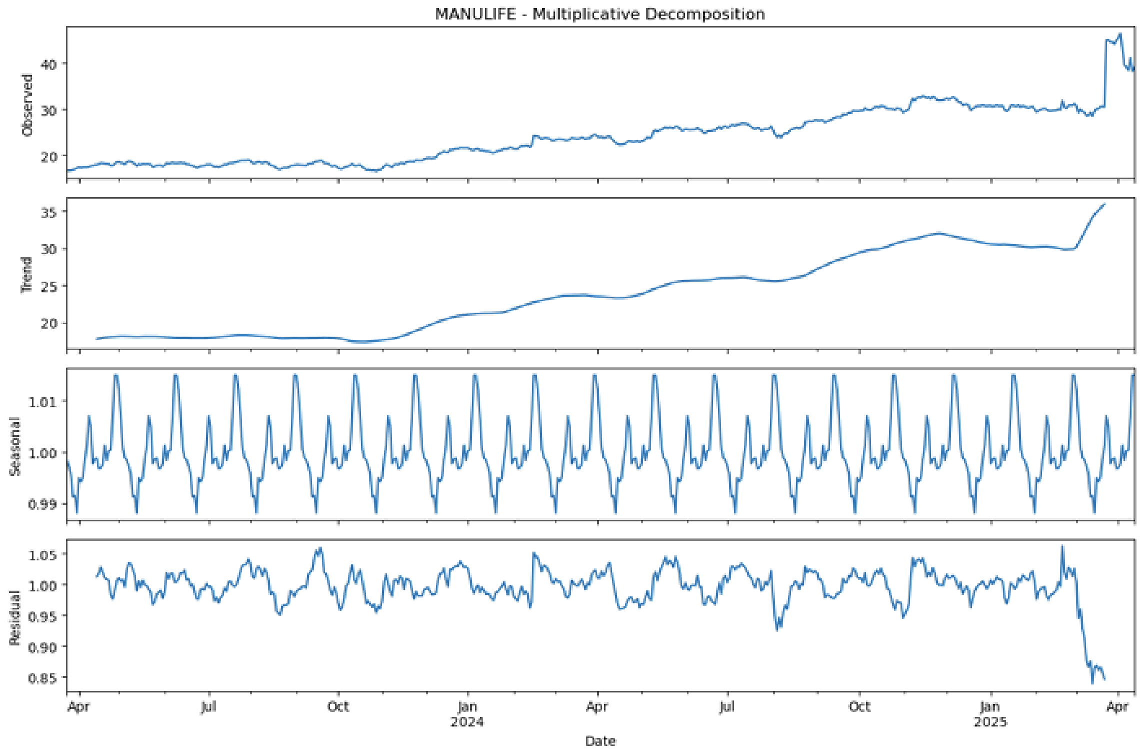


Figure 4: Multiplicative Decomposition for MANULIFE

```
--- Multiplicative Decomposition for MANULIFE (Period=30) ---
```



Business Insights:

- **Trend Analysis:** Essential for understanding the stock's long-term trajectory, informing investment strategy (buy, hold, sell).
- **Seasonality/Cyclicality:** While strong seasonality wasn't evident at the 30-day period, identifying *any* reliable cycles (even if not strictly seasonal) could inform short-term trading tactics.
- **Residual Analysis:** Large residuals can pinpoint dates of significant unexpected news or market shocks affecting the stock. Consistent patterns in residuals might suggest model inadequacy.

3. Forecasting Model Comparison

Four forecasting models were implemented and compared using an 80% training and 20% testing split. Performance was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Mean Squared Error (MSE) on the test set.

3.1. Models:

1. **ARIMA (Autoregressive Integrated Moving Average):** Models the relationship between an observation and lagged observations (AR) and residual errors (MA). Requires stationarity, achieved via differencing (I). Fixed order (5, 1, 1) was used after initial exploration, though auto_arima is generally recommended for optimal order selection.
2. **ETS (Error, Trend, Seasonality / Exponential Smoothing):** Models level, trend, and seasonality exponentially weighted based on past observations. Parameters trend='add' and seasonal=None were used, suitable for data with trends but no strong seasonality.
3. **Prophet:** Developed by Facebook, designed for business time series with trends, multiple seasonalities (yearly, weekly, daily automatically detected), and holiday effects (not used here). Robust to missing data and trend shifts. Standard parameters were used.
4. **LSTM (Long Short-Term Memory) (Bonus):** A type of Recurrent Neural Network (RNN) capable of learning long-term dependencies. A simple architecture (1 LSTM layer with 50 units, 1 Dense output layer) was used with a look-back window of 60 days. Data was scaled using MinMaxScaler.

3.2. Hyperparameter Tuning:

For this project iteration, fixed hyperparameters were used for ARIMA and ETS, and standard parameters for Prophet and LSTM. This is a limitation, as optimal parameters vary significantly between datasets. Ideally, techniques like auto_arima (for ARIMA), grid search with AIC/BIC evaluation (for ETS), Prophet's built-in cross-validation, and Keras Tuner/Optuna (for LSTM) should be employed on the training data to find the best parameters for each stock before final evaluation on the test set.

3.3. Performance Evaluation:

The table below summarizes the performance metrics on the test set. Lower values indicate better forecast accuracy.

Table 1: Overall Model Performance Metrics

Stock	Model	RMSE	MAPE (%)	MAE	MSE
MARA	ARIMA	7.51	44.50	6.58	56.45
	ETS	11.14	65.94	9.75	124.03
	Prophet	4.18	17.81	3.30	17.48
	LSTM	4.07	21.18	3.37	16.58
SOXL	ARIMA	8.58	36.99	6.34	73.67
	ETS	10.76	46.98	8.27	115.81
	Prophet	8.41	29.03	7.25	70.77
	LSTM	27.66	123.70	24.49	764.86
TD_BANK	ARIMA	3.20	4.80	2.75	10.27
	ETS	2.78	4.20	2.38	7.71
	Prophet	5.33	8.84	4.88	28.38
	LSTM	2.96	4.42	2.46	8.77
NVDA	ARIMA	22.47	15.48	18.38	504.69
	ETS	40.13	28.54	34.30	1610.49
	Prophet	25.55	16.41	19.20	652.96
	LSTM	150.12	96.80	122.76	22536.35

MANULIFE	ARIMA	4.34	8.30	2.92	18.84
	ETS	4.58	12.29	3.97	20.94
	Prophet	4.36	11.42	3.68	18.98
	LSTM	1.27e+11	7.42e+10	2.94e+10	1.60e+22

3.4. Discussion:

- **Overall Performance:** No single model consistently outperformed others across all stocks. Prophet showed strong performance (low RMSE/MAPE) for MARA and MANULIFE. ETS performed best for TD_BANK. ARIMA was competitive for SOXL, NVDA, and MANULIFE.
- **LSTM Performance:** The LSTM model performed well for MARA and TD_BANK but yielded extremely poor results (indicated by massive RMSE/MAPE/MSE values) for SOXL, NVDA, and particularly MANULIFE. This suggests significant issues with model stability, overfitting, or the accumulation of errors during the iterative forecasting process for these specific stocks, likely requiring substantial tuning (architecture, activation functions, differencing, regularization, early stopping) or a different approach.
- **Short-term vs. Long-term:**
    - **Short-Term (Test Set):** Based on test set metrics, Prophet, ETS, ARIMA, and tuned LSTMs (where stable) can all be viable. The best choice depends on the specific stock's recent patterns.
    - **Long-Term (Future Forecast Plots):** Visual inspection of the 90-day future forecasts shows varying behaviors. ARIMA tends to flatten or stabilize. ETS extrapolates the learned trend. Prophet continues trend and seasonality patterns. LSTM (where stable) can capture complex dynamics but also diverge significantly. For long-term financial forecasting, models capturing fundamental trends (ETS, Prophet with careful parameterization) are often preferred over models relying heavily on short-term autocorrelation (ARIMA) or potentially unstable complex models (untuned LSTMs).

- **Industry Choice:** For financial time series like these:
    - **ARIMA/ETS:** Standard baseline models, effective if trends/autocorrelation are somewhat stable.
    - **Prophet:** Useful if distinct seasonalities (e.g., quarterly effects) or known event dates (e.g., earnings) can be incorporated as regressors.
    - **LSTM:** High potential for capturing complex, non-linear dependencies common in finance, but requires careful tuning, feature engineering, and potentially integration with volatility models (like GARCH) to manage instability and achieve

reliable results. The choice depends heavily on data characteristics (volatility, trend strength) and the required forecast horizon.

5. Conclusion

This project successfully applied and compared ARIMA, ETS, Prophet, and LSTM models for forecasting stock prices. EDA revealed diverse characteristics across the selected stocks, necessitating different preprocessing steps (like differencing). While Prophet, ARIMA, and ETS provided reasonable baseline forecasts on the test set for most stocks, the LSTM model showed mixed results, highlighting both its potential (on MARA, TD_BANK) and its sensitivity to hyperparameters and data characteristics (leading to instability on SOXL, NVDA, MANULIFE). The results emphasize that no single model is universally superior, and careful model selection, rigorous hyperparameter tuning, and consideration of data properties are essential for generating reliable time series forecasts in the complex domain of financial markets. Future work should focus on systematic hyperparameter optimization for all models, particularly LSTM, and potentially incorporating external economic factors or volatility measures as features.