📊 Stock Price Prediction Report

📌 Introduction

This report provides an overview of the stock price prediction model, detailing data preprocessing, feature engineering, model selection, and performance evaluation.

📌 Data Preprocessing

- Converted `Date` column to datetime format and set it as the index.
- Addressed missing values using interpolation and KNN imputation.
- Applied one-hot encoding to categorical variables (`Ticker`).
- Scaled numerical features using `StandardScaler`.

📌 Feature Engineering

Implemented key financial indicators:

- **Moving Averages (MA5, MA10):** Captures short-term and long-term price trends.
- **Relative Strength Index (RSI):** Measures momentum of price movements.
- **MACD & MACD Signal:** Identifies bullish and bearish trends.
- **Bollinger Bands:** Detects market volatility.
- **Sin-Cos transformation:** Applied to time-based features (day, month) to account for cyclic behavior.

📌 Model Selection & Performance Evaluation

The following models were trained and evaluated:

- **Linear Regression** ✅ (Selected as the final model)
- XGBoost
- LightGBM

- Decision Tree

📌 Model Performance Comparison

Model	MAE	RMSE	R^2 Score
Linear Regression	**3.28**	**4.81**	**0.9971**
XGBoost	6.05	13.99	0.9761
LightGBM	3.92	5.70	0.9960
Decision Tree	3.93	5.99	0.9956

📌 Why Linear Regression?

- **Minimal Overfitting:** The training and test results are closely aligned.
- **High Interpretability:** Feature weights can be analyzed and understood.
- **Best Generalization Performance:** The model consistently performs well on unseen data.

📌 Conclusion & Future Work

The Linear Regression model was chosen due to its balance of accuracy, simplicity, and generalizability. Possible improvements include:

- **Incorporating additional financial features**, such as sentiment analysis from financial news.
- **Experimenting with deep learning techniques**, like LSTMs, for sequential modeling.