COMMODITY PRICE FORECASTING USING MULTI-MARKET FINANCIAL TIME SERIES DATA

GROUP MEMBERS

SAMWEL ONGECHI

MARILYN AKINYI

ERICK MAUTI

LYDIAH CHUMBA

ROSE MIRITI

ISAAC ONYANGO

RODGERS OTIENO

PRESENTED BY: GROUP 5

TM : GEORGE KAMUNDIA

BUSINESS UNDERSTANDING

- Volatility in commodity markets impacts global trade and investments
- Forecasting helps producers, traders, and policymakers make informed decisions

BUSINESS PROBLEM

Predicting commodity returns is particularly challenging due to:

- Market volatility
- Time-lagged effects
- Dependencies across different financial instruments

How can we leverage time series data to predict commodity price movements?

STAKEHOLDERS

- Commodity traders – optimize trading strategies
- Investors –To manage portfolio risk
- Policy makers and regulators – anticipate inflationary trends
- Financial Analysts – provide data-driven forecasts

OBJECTIVES

- To preprocess and integrate multi-market financial time series data for commodities.
- To explore and analyze patterns, correlations, and volatility in commodity returns.
- To develop predictive models for selected commodity prices and spreads.
- To evaluate model accuracy and stability using appropriate performance metrics.
- To generate insights that can support trading strategies and risk management.

DATA UNDERSTANDING

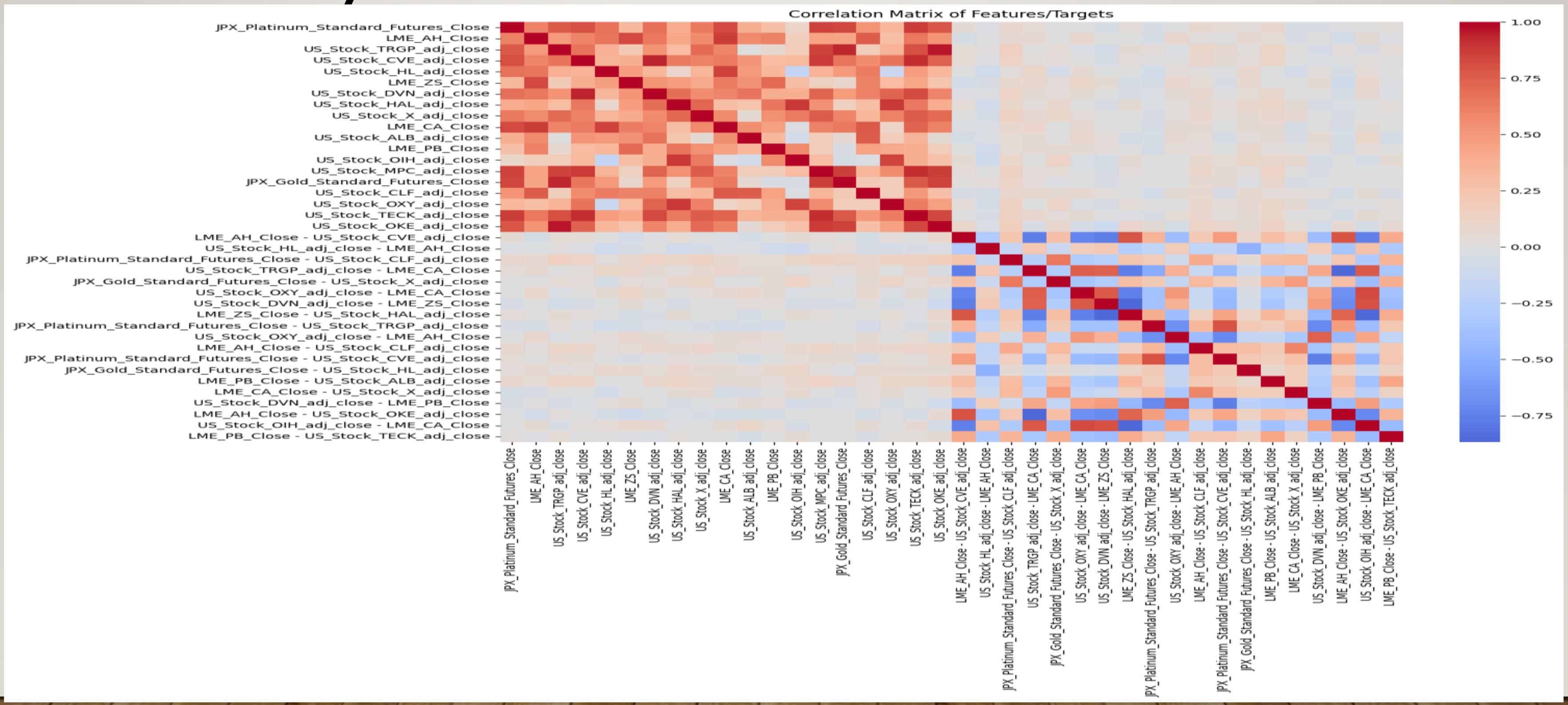
- Sources: Kaggle (London Metal Exchange (LME), Japan Exchange Group (JPX), U.S. stock markets.
- Target pairs and Train labels
- **Time horizon**: Daily
- **Challenges**: missing values, seasonality, volatility, external shocks

VISUALIZATIONS (EDA)

- Understand the distributions of target variables and features.
- Identify trends, patterns, and outliers in the data.
- Explore relationships between features and targets, including correlations and potential predictive signals.
- Visualize time series behaviors since this is a temporal dataset.

VISUALIZATIONS (EDA) CONT…..

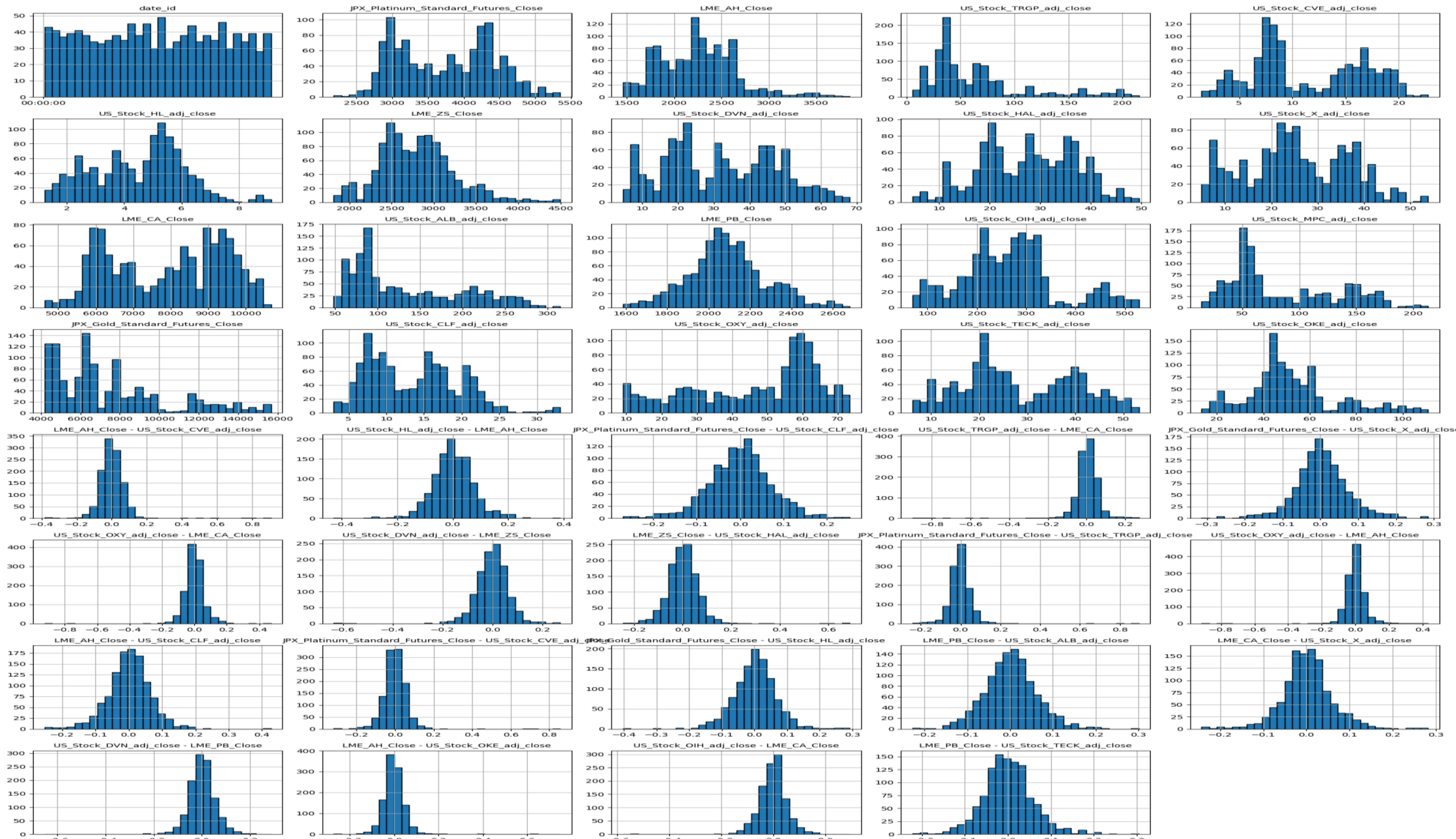
Correlation Analysis



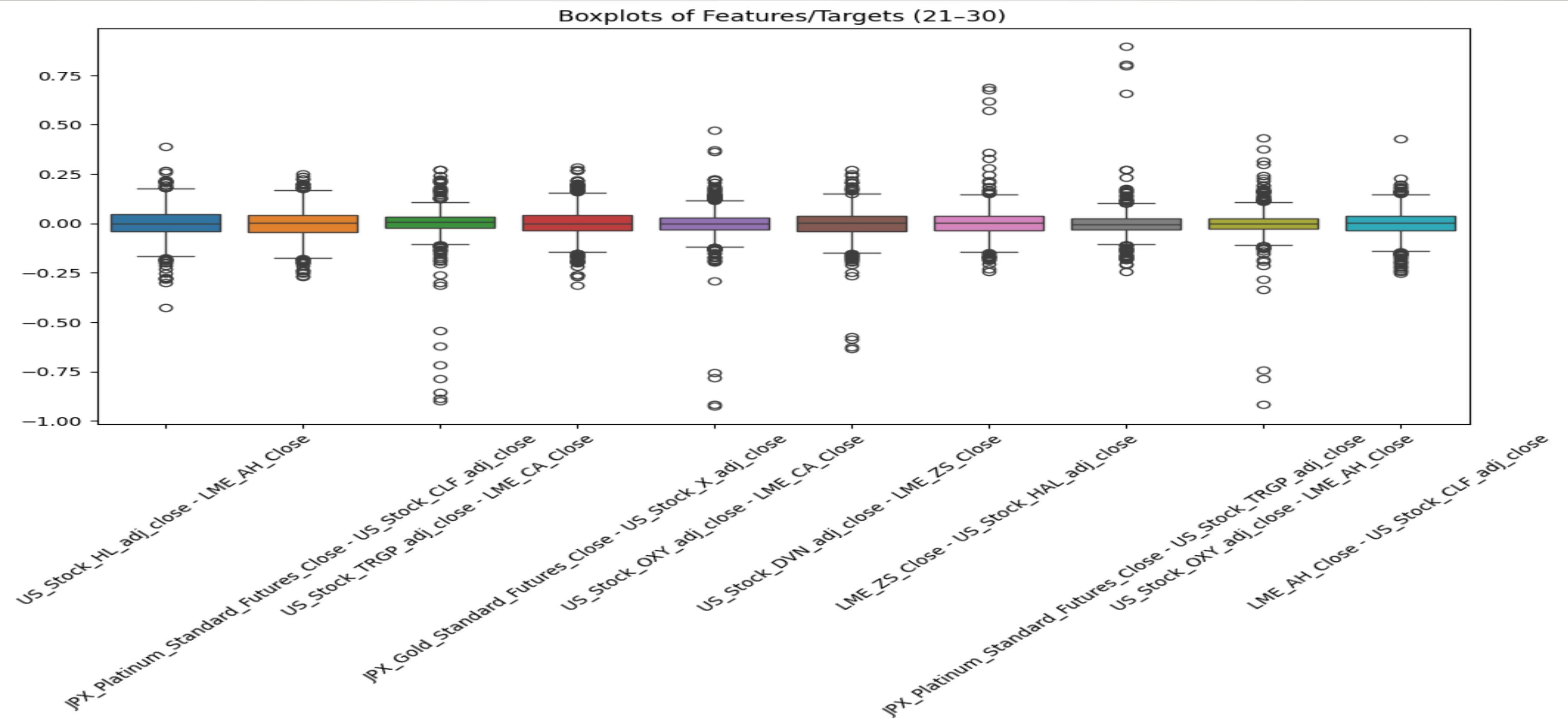
VISUALIZATIONS (EDA) CONT.....

- Applied 0.9 correlation threshold to flag highly correlated pairs.
- A total of 5 pairs exceeded this threshold, meaning they may introduce multicollinearity risks if included together in models.

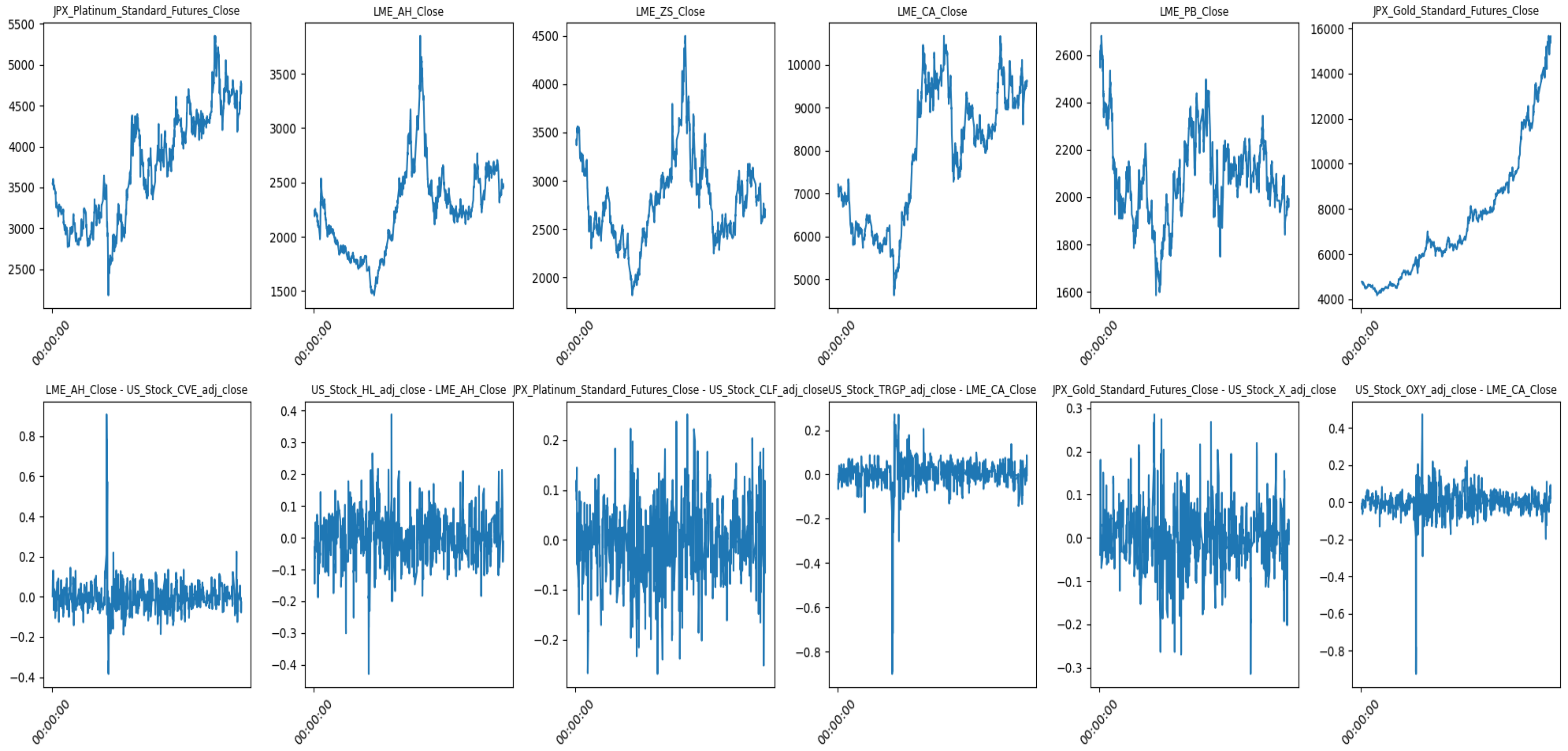
Distribution of Features & Targets



BOXPLOT OF FEATURES & TARGETS



TARGET TRENDS OVER TIME



DATA PREPARATION

- Cleaning missing values & outliers
- Normalization / scaling of variables
- Mapping
- Handling stationarity: differencing, log returns
- Feature engineering: lags, rolling averages, volatility measures

MODELING

BASELINE MODELS

- Linear Regression
- Random Forest
- XGBoost
- LightGBM

ADVANCE MODELS

- Gradient Boosting
- ElasticNet
- SVR
- MLP Regressor

TIME SERIES MODEL- SARIMA

EVALUATION

- Mean Absolute Error (MAE): Average magnitude of prediction errors.
- Root Mean Squared Error (RMSE): Measures error magnitude with greater penalty for larger deviations.
- R-squared (R^2): Proportion of variance explained by the model.
- MAPE (Mean Absolute Percentage Error)
- SMAPE (Symmetric MAPE)
- Directional Accuracy: Percentage of correct predictions in the direction of price movement.

EVALUATION

MAE	RMSE	R2	MAPE	SMAPE	tional Acc	Model	mbinedSco
0.0197	0.0289	0.5504	402.3945	107.2267	0.74	Ridge_Tuned	3.6867
0.0224	0.0332	0.4202	351.0624	108.4613	0.7228	Ridge	3.3827
0.0224	0.0332	0.4194	351.5111	108.4612	0.7227	LinearRegression	3.3807
0.0225	0.0333	0.4154	359.1141	108.7319	0.7222	SARIMA+Ridge	3.3717
0.0232	0.0362	0.3513	346.8279	111.7668	0.7219	RandomForest	3.2334
0.0237	0.0371	0.319	352.7172	111.854	0.7102	GradientBoosting	3.1504
0.024	0.0376	0.3046	375.5492	113.6642	0.7049	LightGBM	3.112
0.0255	0.0397	0.1742	412.7278	115.5255	0.6927	XGBoost	2.8353
0.031	0.0457	-0.041	420.2271	140.6937	0.5804	SVR	2.216
0.0297	0.0451	-0.0058	114.3862	183.1374	0.019	SARIMA	1.5354
0.0297	0.0451	-0.0049	113.0884	185.035	0	ElasticNet	1.5113
0.0297	0.0451	-0.0049	113.0884	185.035	0	Lasso	1.5113
0.0325	0.0474	-0.1467	400.3438	152.7311	0	NaivePersistence	1.2184
0.125	0.1858	-16.9889	2923.832	155.6005	0.5402	MLP	-30.1365

EVALUATION CONT……

MAE	RMSE	R2	MAPE	SMAPE	tional Acc	Model	mbinedSco
0.0197	0.0289	0.5504	402.3945	107.2267	0.74	Ridge_Tuned	3.6867
0.0224	0.0332	0.4202	351.0624	108.4613	0.7228	Ridge	3.3827
0.0224	0.0332	0.4194	351.5111	108.4612	0.7227	LinearRegression	3.3807
0.0225	0.0333	0.4154	359.1141	108.7319	0.7222	SARIMA+Ridge	3.3717
0.0232	0.0362	0.3513	346.8279	111.7668	0.7219	RandomForest	3.2334
0.0237	0.0371	0.319	352.7172	111.854	0.7102	GradientBoosting	3.1504
0.024	0.0376	0.3046	375.5492	113.6642	0.7049	LightGBM	3.112
0.0255	0.0397	0.1742	412.7278	115.5255	0.6927	XGBoost	2.8353
0.031	0.0457	-0.041	420.2271	140.6937	0.5804	SVR	2.216
0.0297	0.0451	-0.0058	114.3862	183.1374	0.019	SARIMA	1.5354
0.0297	0.0451	-0.0049	113.0884	185.035	0	ElasticNet	1.5113
0.0297	0.0451	-0.0049	113.0884	185.035	0	Lasso	1.5113
0.0325	0.0474	-0.1467	400.3438	152.7311	0	NaivePersistence	1.2184
0.125	0.1858	-16.9889	2923.832	155.6005	0.5402	MLP	-30.1365

EVALUATION CONT….

- **Best Model** `Ridge Tuned` achieved the highest Combined Score, outperforming both untuned Ridge and advanced ensemble models.
- -**Strong Baselines:** Ridge` and `LinearRegression` ranked closely behind, showing that linear methods remain competitive.
- -**Hybrid Success:**`SARIMA+Ridge` performed better than standalone SARIMA, proving that combining time-series modeling with machine learning improves results.
- **Tree-based Models:**`RandomForest` and `GradientBoosting` showed strong but slightly weaker performance compared to tuned Ridge.
- **Weaker Models:**`ElasticNet`, `Lasso`, and `NaivePersistence` underperformed significantly, while `MLP` collapsed with highly negative R^2.
- **SARIMA Alone:**Bare SARIMA struggled, but adding Ridge correction significantly boosted its performance.

ENSEMBLE MODELING

Rank	Model	R2	Directional Accuracy
1	Ensemble-Weighted-Top3	0.7227	0.8028
2	Ensemble-Mean-Top3	0.6716	0.784
3	Ridge_Tuned	0.5504	0.74
4	Ridge	0.4202	0.7228
5	LinearRegression	0.4194	0.7227
6	SARIMA+Ridge	0.4154	0.7222
7	RandomForest	0.3513	0.7219
8	GradientBoosting	0.319	0.7102

EVALUATION CONT….

- Best single model R^2: 0.5504
- Best ensemble R^2: 0.7227
- R^2 improvement: +0.1723
- The ensemble outperforms individual models, providing more accurate predictions.

FEATURE IMPORTANCE

Rank	Feature	Importance
1	US_Stock_OIH_adj_close	0.1853
2	US_Stock_HAL_adj_close	0.1216
3	US_Stock_CVE_adj_close	0.1055
4	LME_CA_Close	0.0887
5	US_Stock_OKE_adj_close	0.0882
6	US_Stock_OXY_adj_close	0.0832
7	LME_ZS_Close	0.0813
8	US_Stock_DVN_adj_close	0.0758
9	LME_PB_Close	0.0579
10	US_Stock_TRGP_adj_close	0.0535
11	LME_AH_Close	0.0491
12	US_Stock_TECK_adj_close	0.0279
13	JPX_Platinum_Standard_I	0.0276
14	JPX_Gold_Standard_Futu	0.0157
15	US_Stock_HL_adj_close	0.0132

RECOMMENDATIONS

- **Go with the Ensemble-Weighted-Top3 Model**
- **Energy Sector Features Matter Most**
- **Markets Don't Move in Isolation**

NEXT STEP

- **Strengthen Validation Testing**
- **Smarter Feature Engineering**
- **Build a Monitoring System**

LIMITATIONS

Market regimes change

- The model is tuned for today's conditions but markets evolve
- Over-reliance on energy features might reduce accuracy in tech-led or defensive markets.
- External shocks (wars, pandemics, policy changes)

Validation Gaps

- Back tests are based on past data true future-proofing is untested.
- Assumes markets remain liquid execution could fail during crises.