Stakeholder Report: Stock Price Modeling and Risk Assessment

Executive Summary

This project explored daily stock data for CME Group (CME) using AlphaVantage's API. We engineered predictive features, developed regression and classification models, and evaluated them under uncertainty. The results highlight key assumptions, risks, and limitations, with implications for future improvements.

1. Data Collection and Preparation

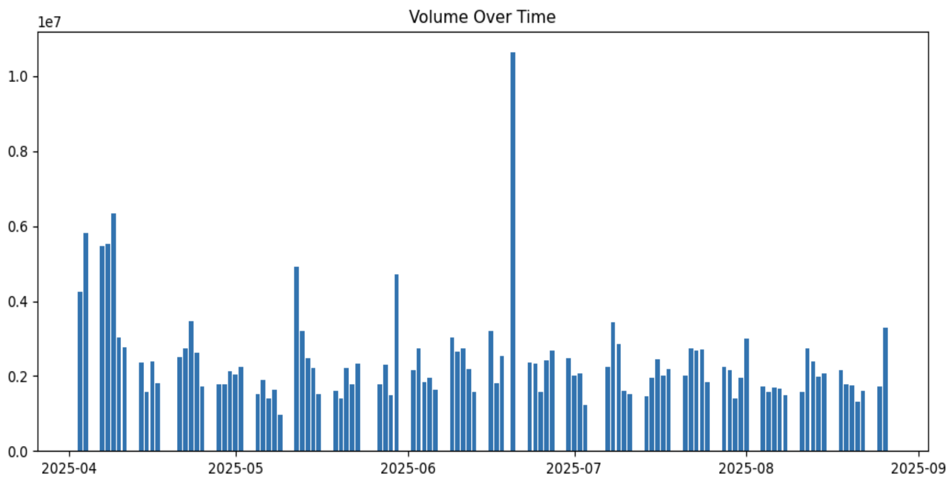
- Source: AlphaVantage API (TIME_SERIES_DAILY)
- Period: Last ~100 trading days (April–August 2025)
- Features Collected: open, high, low, close, volume
- Feature Engineering:
    - Daily returns (pct_change)
    - Rolling averages (5-day moving mean of close, volume)
    - Lagged variables (1-day lag close/returns)
    - Price range (high - low), opening/closing difference (close - open)

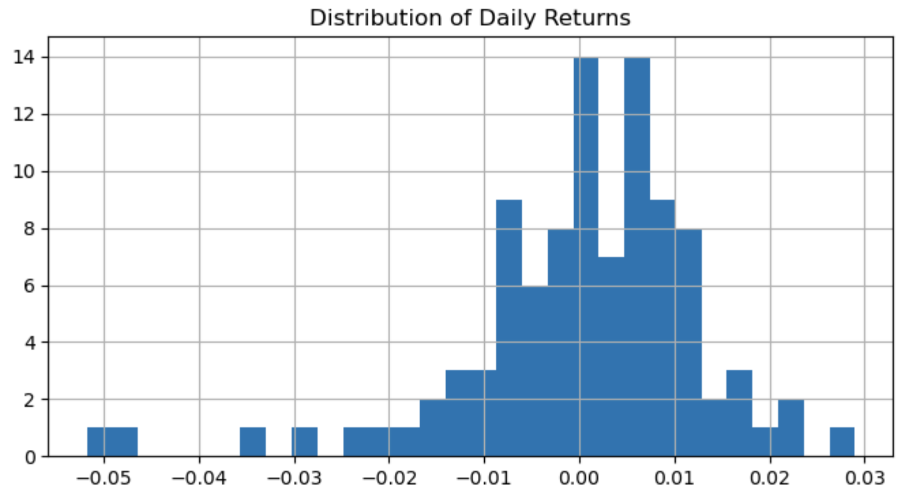
2. Exploratory Data Analysis (EDA)

- Closing prices fluctuated between $253 and $290, showing short-term volatility.
- Returns distribution was centered near zero, with mild negative skew.
- Volume distribution was highly skewed, suggesting occasional abnormal trading days.
- Boxplots showed typical price clustering, but with several outliers in both high and low prices.

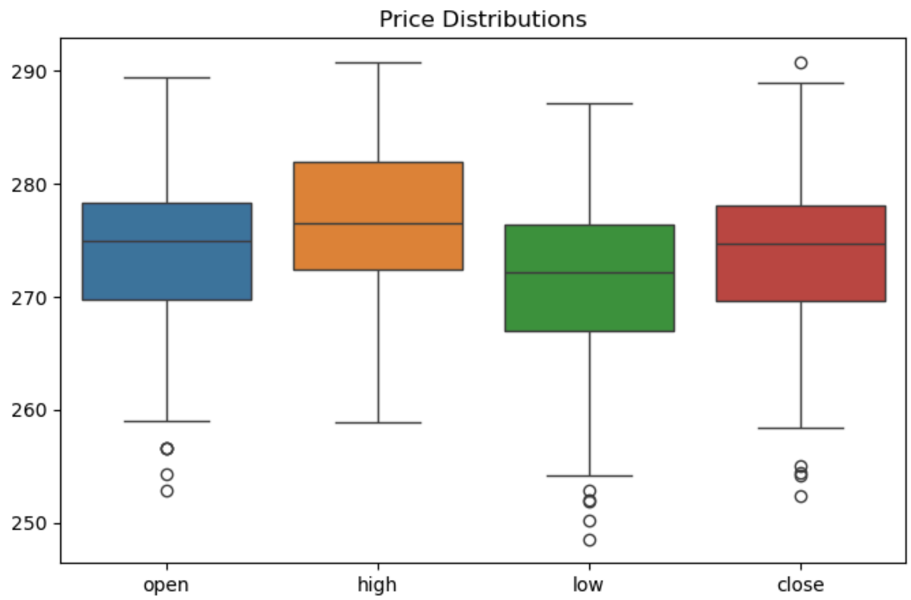
Takeaway: Data reflects a relatively stable price channel but includes volatility spikes, consistent with financial time series behavior.


Closing Price Over Time
Close
Price
290
285
280
275
270
265
260
255
2025-04
2025-05
2025-06
2025-07
2025-08
2025-09
Date


Volume Over Time
1e7
1.0
0.8
0.6
0.4
0.2
0.0
2025-04
2025-05
2025-06
2025-07
2025-08
2025-09




```
: Text(0.5, 1.0, 'Price Distributions')
```



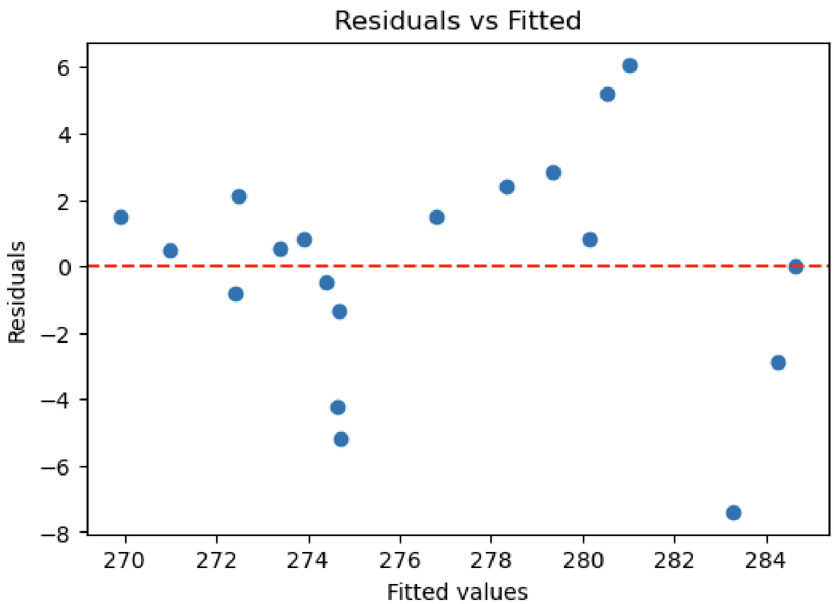
3. Modeling Approaches

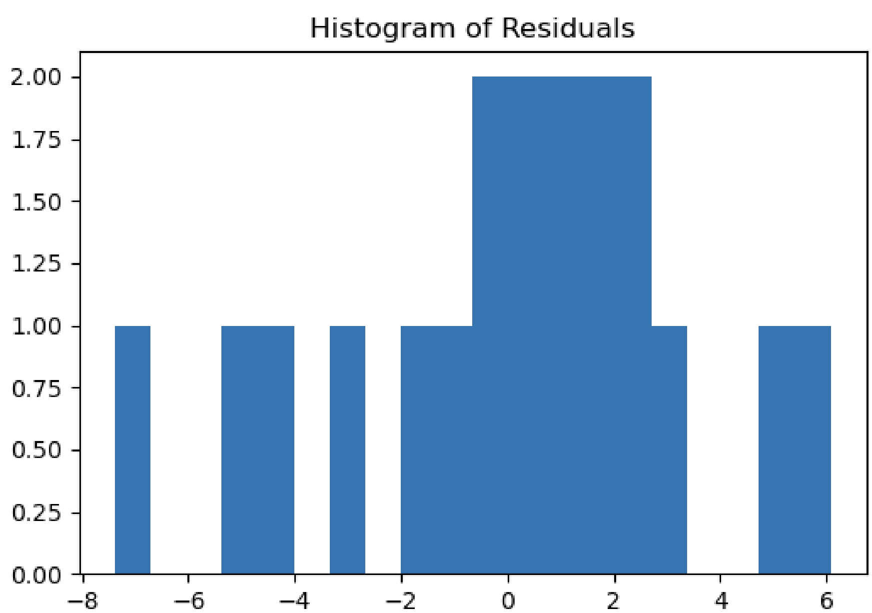
Regression Track

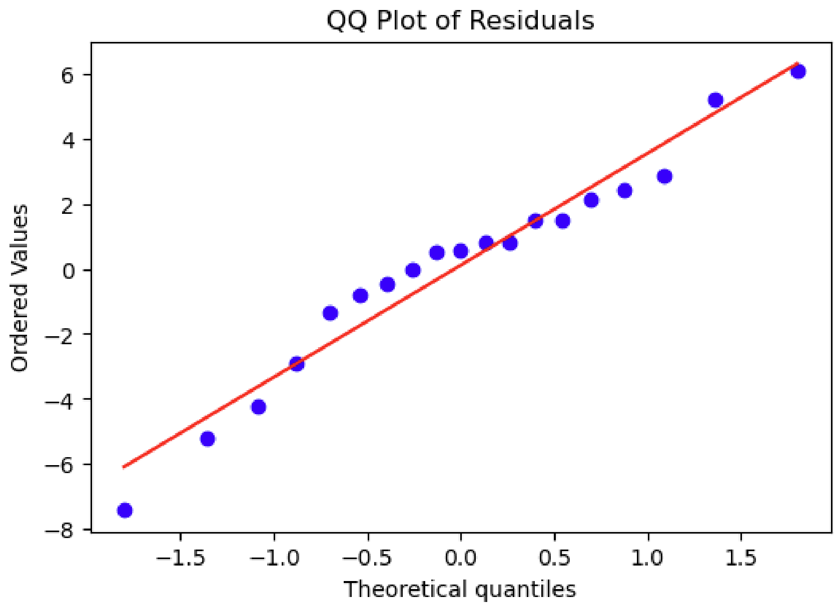
- **Goal:** Predict next-day closing price.
- **Model:** Linear Regression with engineered features.
- **Performance:**
    - R2= 0.636 (reasonable explanatory power on training data)
    - RMSE ≈ **3.24**, indicating modest absolute prediction error.
- **Residuals:** Slight non-linearity present; QQ plot showed near-normal errors, though extreme values deviate.

Classification Track

- Goal: Predict whether next-day return > 0 (up vs. down).
- Model: Logistic Regression with scaling.
- Performance:
    - Accuracy ≈ 44%
    - Recall for “up” days better than “down” days, indicating model bias toward positive returns.
- Interpretation: Poor predictive performance, near random, suggesting engineered features did not capture return direction well.



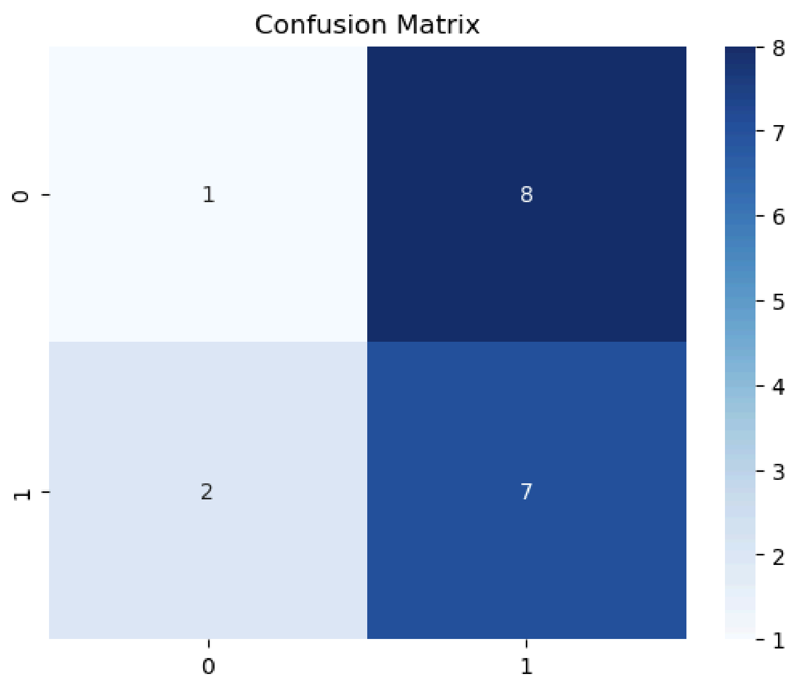





```
              precision    recall  f1-score   support

           0       0.33      0.11      0.17         9
           1       0.47      0.78      0.58         9

    accuracy                           0.44        18
   macro avg       0.40      0.44      0.38        18
weighted avg       0.40      0.44      0.38        18
```



4. Uncertainty and Sensitivity Analysis

Confidence Intervals

- Parametric CI: Gaussian-based CI around regression fit showed narrow uncertainty bands.
- Bootstrap CI: Wider intervals, reflecting sampling variability.

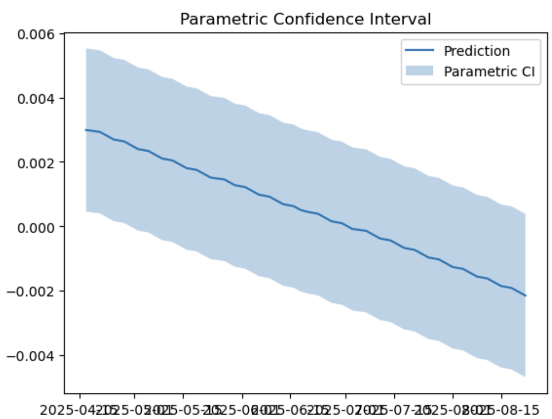
Scenario Comparisons

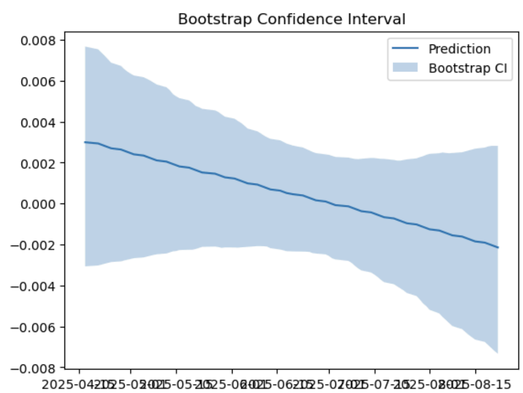
- Mean vs. Median Imputation: Nearly identical RMSE (~0.0120) and $R2$ (~0.015).
- Drop Missing: Performance degraded (RMSE = 0.0206).

Takeaway: Model is robust to imputation method but performs poorly when missing values are dropped.

Subgroup Diagnostics

- Segmented by volatility: Residuals varied, with higher variance in high-volume regimes.
- Takeaway: Model stability decreases under volatile trading conditions.





	mean	median
rmse	0.012037	0.012037
r2	0.015278	0.015278

5. Key Assumptions

- Linear relationship between features and targets.
- Stationary data generating process over the observed window.
- Missing values are random (imputation does not bias outcomes).

6. Risks

- Structural breaks: Regime changes in financial markets may render features obsolete.
- Linearity: Financial data often exhibit non-linear dynamics (e.g., jumps, volatility clustering).

- Data limitations: Limited time horizon (~100 days) reduces generalizability.

7. Recommendations and Next Steps

Modeling Improvements:
- Introduce non-linear models (e.g., Random Forest, Gradient Boosting, or LSTMs).
- Add volatility indicators (e.g., rolling variance, GARCH).

Extend dataset to capture multiple years and market conditions.

Risk Communication:
- Current results cannot reliably forecast returns, but provide baseline understanding of feature relationships.
- Decision-makers should not rely on this model for trading but may use it for exploratory research.

Future Deliverables:
- Interactive dashboard to update forecasts with new data.
- Additional scenario analysis under stress (e.g., simulated market shocks).

Conclusion

This project establishes a baseline pipeline for stock data ingestion, feature engineering, and predictive modeling. The regression model shows partial predictive power, while classification struggles to distinguish return direction. Confidence interval and scenario analysis underscore the model's limitations and sensitivity to assumptions.

For stakeholders, the key insight is that **linear models alone are insufficient for reliable return prediction**, and that **non-linear, volatility-aware models** should be prioritized moving forward.