Xyphramin Technologies Assessment - Volatility vs Volume

1 Problem Statement Understanding

The task is to investigate the mathematical relationship between intraday volatility and volume of financial instruments using historical data of ETFs and stocks. The goal is to:

- Extract features such as volatility and volume deltas.
- Discover a reliable mathematical equation that models their relationship.
- Evaluate the robustness of the equation across various assets.
- Submit the analysis in a clear, modular, and reproducible format.

2 Approach

2.1 Data Collection & Preparation

We were provided historical CSV files for:

- NVDA (Stock)
- QQQ (ETF)
- SOXS (Inverse ETF)

Each dataset was:

- Sorted chronologically.
- Cleaned for missing values.
- Parsed to compute key features.

2.2 Feature Engineering

We derived two key engineered features:

Intraday Volatility:

$$\text{Volatility}_t = \frac{\text{High}_t - \text{Low}_t}{\text{Close}_t}$$

Deltas:

$$\Delta\text{Volatility}_t = \text{Volatility}_t - \text{Volatility}_{t-1}$$

$$\Delta\text{Volume}_t = \text{Volume}_t - \text{Volume}_{t-1}$$

2.3 Modeling

We hypothesized that:

$$\Delta\text{Volatility}_t = f(\Delta\text{Volume}_t)$$

Models Used:

- Linear Regression
- Polynomial Regression (Degree 2)

3 Mathematical Explanation

3.1 Linear Model

$$\Delta V_t = a \cdot \Delta\text{Volume}_t + b$$

This is a simple baseline model. However, it cannot capture nonlinear market reactions.

3.2 Polynomial Model (Quadratic)

$$\Delta V_t = a(\Delta\text{Volume}_t)^2 + b(\Delta\text{Volume}_t) + c$$

This form:

- Captures nonlinear relationships.
- Allows for asymmetrical response to volume spikes.
- Balances interpretability and flexibility.

Why Degree-2 Polynomial?

- Volume spikes often cause disproportionate changes in volatility.
- Markets exhibit non-linear behavior under stress.
- Quadratic functions capture turning points effectively.

4 Model Results

Each model was fitted, and coefficients & $\mathbf{R^2}$ scores saved. Here are the final equations:

4.1 NVDA (Stock)

$$\Delta V = -1.75 \times 10^{-19}(\Delta\text{Volume})^2 + 3.89 \times 10^{-10}(\Delta\text{Volume}) + 0.00013$$
$$R^2 = 0.288$$

4.2 QQQ (ETF)

$$\Delta V = -1.31 \times 10^{-19}(\Delta\text{Volume})^2 + 1.76 \times 10^{-10}(\Delta\text{Volume}) + 0.000095$$
$$R^2 = 0.278$$

4.3 SOXS (Inverse ETF)

$$\Delta V = 3.16 \times 10^{-18}(\Delta\text{Volume})^2 + 1.64 \times 10^{-9}(\Delta\text{Volume}) - 0.000098$$
$$R^2 = 0.051$$

5 Intuition Behind Model Selection

- Volume is a key signal for market activity and risk.
- Nonlinear models (especially degree-2 polynomials) better reflect market dynamics.
- Quadratic models are easy to fit, interpret, and visualize.

6 Project Outcome & Review

6.1 Achievements

- Modular codebase with reusable components.
- Jupyter notebook for interactive analysis.
- Model equations and visuals saved to `results/` and `plots/`.

6.2 Deliverables

- `plots/`: Graphs for fitted models.
- `results/model_equations.txt`: Saved equations.
- `README.md`: Full project summary.

6.3 Review

- NVDA & QQQ showed consistent moderate correlation.
- SOXS had weak correlation, likely due to leveraged/inverse structure.
- Polynomial regression provided a simple yet effective solution.

7 Summary

This project demonstrated:

- How engineered features like volatility and volume deltas help model price behavior.
- That quadratic equations balance simplicity and explanatory power.
- A robust, modular approach for exploratory financial data science.