Macro-Linked Energy ETF

Checkpoint B – Research Progress Report
Author: Vrishani Shah
Course: MSDS 413 – Time Series and Forecasting
Instructor: Tom Miller
Date: November 2025
Repository: github.com/vrishanishah20/macro-linked-energy-etf

1. Project Overview

This project develops a **macro-linked Energy ETF** that dynamically adjusts exposure to major U.S. energy equities - **XOM**, **CVX**, and **XLE** based on macroeconomic and commodity-market indicators.
Checkpoint A established the conceptual foundation; Checkpoint B demonstrates progress in **data acquisition, exploratory analysis, and initial model prototyping**.
Research Question:
Can macroeconomic and commodity-linked signals improve the risk-adjusted performance of an energy-sector ETF relative to a passive benchmark such as XLE or the S&P 500?

2. Progress Since Checkpoint A

- Implemented initial data-pull scripts using `yfinance` and `pandas_datareader` for daily prices (XOM, CVX, XLE, WTI `CL=F`).
- Retrieved macroeconomic series—Fed Funds Rate, CPI, Industrial Production from FRED API; merged on date index.
- Imported weekly crude-inventory data from EIA API and forward-filled to daily frequency.
- Generated daily log-return series and aligned trading days.
- Ran ADF tests → prices non-stationary ($p > 0.1$); returns stationary ($p < 0.01$).
- Produced rolling 60-day correlation plots between WTI and XLE returns.
- Added event dummies for OPEC announcements, COVID-19 shock, and Russia-Ukraine conflict.

3. Exploratory Data Analysis

3.1 Summary Statistics (2015 – 2024)

Series	Mean (%)	Std Dev (%)	ADF p-value	Comment
WTI returns	0.03	2.15	0.00	Stationary
XLE returns	0.04	1.58	0.00	Stationary
XOM returns	0.05	1.42	0.00	Stationary
CVX returns	0.04	1.37	0.00	Stationary

3.2 Pairwise Correlations

Relationship	ρ	Economic Interpretation
WTI ↔ XLE returns	0.61	Energy sector tracks oil prices
WTI ↔ XOM returns	0.59	Direct commodity exposure
Fed Funds Rate ↔ XLE	–0.24	Tight policy reduces sector returns
Inventory Δ ↔ WTI	–0.31	Inventory builds lower prices

3.3 Visual Insights

- **Figure 1.** Rolling 60-day correlation between WTI and XLE returns (peaks ≈ 0.8 in 2022 oil surge).
- **Figure 2.** ACF/PACF of XLE returns → low autocorrelation beyond lag 1 → ARIMA(1,0,1) baseline.
- **Figure 3.** Event markers show oil-volatility spikes around OPEC and geopolitical news.

4. Model Preparation and Early Results

4.1 Baseline Specification

$$r_{t+1}^{XLE} = \alpha + \beta_1 r_t^{WTI} + \beta_2 \Delta \text{FFR}_t + \beta_3 \Delta \text{Inv}_t + \varepsilon_{t+1}$$

4.2 Preliminary ARIMAX Prototype

Variable	Coefficient	t-stat	p-value
Intercept	0.02	1.9	0.06
WTI Return	0.43	5.1	0.00 ***
Δ FFR	–0.12	–2.3	0.02 **
Δ Inventory	–0.07	–1.8	0.08 *

Model $R^2 \approx 0.18$; residuals uncorrelated (Q-test $p > 0.2$).
Interpretation: Energy returns rise with oil shocks and fall with monetary tightening and inventory builds.

4.3 Planned Extension

- Upgrade to VAR(2) with endogenous series (XOM, CVX, WTI).
- Compute Impulse-Response Functions for shock propagation.
- Derive macro-signal-based ETF re-weighting rule.

5. ETF Signal Rule Design

$$C_t = w_1 \cdot z(r_t^{WTI}) + w_2 \cdot z(\Delta \text{Inv}_t) - w_3 \cdot z(\Delta \text{FFR}_t)$$

Allocation Policy:

- If $C_t > 0.75$ → Overweight Energy (+20 %)
- If $C_t < -0.75$ → Underweight Energy (–20 %)
- Else → Neutral
  Backtesting scheduled for Checkpoint C.

6. Next Steps (Toward Checkpoint C & Final)

1. Finalize data pipeline in Python (/src).
2. Fit VAR/ARIMAX models with rolling out-of-sample windows.
3. Backtest ETF signal and benchmark vs XLE.
4. Report CAGR, Sharpe, and Max Drawdown.
5. Integrate results into final PDF research report.

7. References

- Bouchouev, I. (2023). Virtual Barrels: Quantitative Trading in the Oil Market. Springer.
- Downey, M. (2009). *Oil 101.* Wooden Table Press.
- Edwards, D. W. (2017). Energy Trading & Investing (2nd ed.). McGraw-Hill.
- Hamilton, J. D. (2009). "Causes and Consequences of the Oil Shock of 2007–08." *Brookings Papers on Economic Activity.*
- Grinold, R., & Kahn, R. (2000). *Active Portfolio Management.* McGraw-Hill.
- López de Prado, M. (2018). Advances in Financial Machine Learning. Wiley.
- Markowitz, H. (1952). "Portfolio Selection." *Journal of Finance.*
- Sharpe, W. F. (1963). "Capital Asset Prices." *Journal of Finance.*