Annotated Workflow: Single-Stage Relative Sector Rotation Model

Objective: To develop and validate a quantitative model that ranks the 11 GICS sectors based on predicted relative performance. The project will conduct an A/B test between two distinct feature sets based on specific economic drivers to determine the most robust approach. The final model will be used to construct a monthly rotating portfolio that aims to outperform benchmarks while managing regime risk.

Phase 1: Data Acquisition & Preparation

Goal: To assemble a complete, clean, and historically consistent daily dataset for all assets and macro indicators, ensuring all data is point-in-time correct.

1.1 Primary Data Sources

This section outlines the required data inputs for the model and strategy.

- **Sector ETFs (11 GICS Sectors):** These are the 11 tradable assets in the investment universe. Each Exchange-Traded Fund (ETF) is a basket of stocks representing a specific sector of the economy (e.g., Technology, Healthcare). The model's primary objective is to rank these 11 assets based on predicted performance.
- **Benchmark Data:** These assets serve as yardsticks for performance measurement and as defensive holdings.
    - SPY: An ETF tracking the S&P 500 Index, representing the broad US stock market. It is used as a "weather vane" to determine the overall market's health for the risk management framework.
    - SHY: An ETF holding short-term US government bonds, considered a low-risk "cash equivalent." The strategy allocates to this asset during dangerous market conditions identified by the risk management rules.
- **Macro Data (FRED & other sources):** This is macroeconomic data providing context on the state of the broader economy.
    - DGS10 (10-Year Treasury Yield): The interest rate on 10-year government bonds, a key driver for rate-sensitive sectors.
    - BAMLH0A0HYM2 (High-Yield Credit Spread): The spread between high-yield corporate bonds and risk-free government bonds. It is a sensitive measure of financial stress and investor risk appetite.
    - WTICL (WTI Crude Oil Price): A direct economic driver for the Energy sector (XLE).

1.2 Historical Data Challenges & Solutions

This section addresses a critical real-world data problem: two of the sector ETFs, Real Estate (XLRE) and Communication Services (XLC), have significantly shorter histories than the other nine.

- **Limited History:** A direct backtest to the year 2000 is not possible, as the data for XLRE only begins in 2015 and for XLC in 2018.
- **Backfilling Solution:** To enable a longer historical simulation, a "proxy" or "synthetic" history is created for these two ETFs for the period before their inception.
- **The "Bootstrap" Compromise:** The plan makes a crucial distinction between backtesting and model training. The **model itself will only be trained on real, traded ETF data** whenever possible to ensure it learns from actual market behavior. However, to enable a backtest period starting in 2019, the very first model must be trained on data from 2016-2018, which necessarily includes the synthetic proxies. As the backtest's expanding window moves forward in time, this synthetic data is quickly replaced by real data in the training set, making the model progressively more robust.

Phase 2: Feature Engineering

Goal: To create and A/B test two distinct, disciplined feature sets based on targeted economic drivers.

2.1 Feature Set A: Economic Driver Model (9 Features)

This set is based on specific economic drivers rather than generic macro factors.

- **Sector-Specific (5):**
  1. return_63d (Medium-term momentum)
  2. return_252d (Long-term trend)
  3. volatility_63d (Risk measure)
  4. sharpe_63d (Risk-adjusted return)
  5. rank_return_63d (Cross-sectional relative strength)
- Targeted Macro & Regime (4):
  6. spy_above_200d_ma (Non-negotiable regime filter)
  7. oil_price_momentum_63d (Specific driver for Energy sector)
  8. treasury_10y_level (Absolute level of interest rates, driver for Financials/Utilities/REITs)
  9. hy_credit_spread (A sensitive measure of credit risk and risk appetite)

2.2 Feature Set B: Economic Driver + Rate Momentum Model (10 Features)

This set tests if adding the momentum of interest rates provides additional predictive

power. It includes all 9 features from Set A, plus one additional feature.

- **All Features from Set A (9)**
- Additional Feature (1):
  10. treasury_10y_1m_change (Recent change/momentum in interest rates)

2.3 Feature Engineering Rules

- **Data Quality:** A minimum of 252 days of history is required for feature generation; missing values will be forward-filled (max 5 days); outliers will be capped.
- **Cross-Sectional Calculations:** Rankings and other cross-sectional features will be calculated only on the set of available sectors at each point in time.

Phase 3: Target Engineering & Timing

Goal: To define the model's prediction target and enforce a realistic, bias-free timeline.

- **3.1 Target Variable:** The target is return_21d_forward, the log return of a sector over the next 21 trading days. The goal is to accurately **rank** sectors, not predict the exact return value. The model is successful if it consistently identifies the future leaders and laggards.
- **3.2 Prediction-Implementation Timing:** A strict T-2 timeline will be used to prevent look-ahead bias. Predictions are generated two days before month-end, and trades are executed on the last trading day of the month.
- **3.3 Walk-Forward Validation:** The models will be tested using an expanding window, retrained monthly, with a minimum initial training period of 36 months. This ensures the model is always tested on unseen data. The primary out-of-sample test period is 2019-2024.

Phase 4: Model Development & Comparison

Goal: To conduct a rigorous A/B test of the two primary feature sets and model complexities.

4.1 Model Comparison Framework

The core of the project is to compare the out-of-sample performance of four primary model-feature combinations against two baselines to determine if complexity adds value.

- **Model/Feature Combinations:**
  1. XGBoost with Feature Set A (9 features)
  2. Elastic Net with Feature Set A (9 features)

3. XGBoost with Feature Set B (10 features)
4. Elastic Net with Feature Set B (10 features)

- **Baseline Models:**
  1. Simple momentum (return_63d ranking only)
  2. Equal-weight all sectors

The combination with the best risk-adjusted performance in the walk-forward validation will be selected as the final strategy.

4.2 & 4.3 Feature Selection & Hyperparameter Optimization

- **Feature Selection:** For the Elastic Net models, which are sensitive to multicollinearity, feature selection will be performed using Mutual Information and VIF analysis to create a decorrelated feature subset.
- **Hyperparameter Optimization:** A Grid Search will be used to tune the internal settings (hyperparameters) for both XGBoost and Elastic Net models to find their optimal configurations.

Phase 5: Risk Management Framework

Goal: To protect capital. This phase defines the safety rules that override the model's trading signals.

- **5.1 Primary Controls:** These are the main on/off switches for the strategy.
  - **Market Regime Filter:** If the S&P 500 is in a major downtrend (below its 200-day moving average), the model's rankings are ignored, and the portfolio is moved to the safety of cash/bonds.
  - **Performance Stop-Loss:** If the strategy itself underperforms its benchmark by a significant amount over a long period, it is assumed to be "broken," and trading is halted.
- **5.2 Secondary Controls:** These are "dimmer switches" that reduce risk without shutting down entirely. They react to rising market volatility or stress in the credit markets by proactively reducing the amount of capital invested.
- **5.3 Portfolio Construction Rules:** This section defines the two clear states of the portfolio. In a "Normal Market," the model's top 3 ranked sectors are held. In a "Risk-Off" state (triggered by the primary controls), the portfolio holds only defensive cash/bond assets.

Phase 6: Backtesting & Evaluation

Goal: To run the historical simulation and grade the strategy's performance using fair and professional metrics.

- **6.1 Performance Metrics:** The evaluation focuses on professional, risk-adjusted

metrics, not just raw returns.
    - **Sharpe Ratio:** The classic measure of return achieved per unit of risk taken.
    - **Sortino Ratio:** An enhancement of the Sharpe Ratio that only penalizes for "bad" (downside) volatility.
    - **Max Drawdown & Calmar Ratio:** These measure the largest peak-to-trough loss the portfolio experienced and the return relative to that loss.
- **6.2 Robustness Testing:** The strategy is stress-tested by slightly changing its rules (e.g., using a different moving average) to ensure the original success was not due to luck. The simulation also zooms in on specific historical crises to analyze the strategy's behavior under stress.
- **6.3 Model Diagnostics:** This involves looking "under the hood" of the model to analyze which features were most important for its predictions and to measure the stability of its rankings over time.

Phase 7: Implementation & Monitoring

Goal: To define the real-world process for running the strategy and establishing a plan for when to stop trading it.

- **7.1 Production Pipeline:** A disciplined, repeatable monthly checklist to ensure the strategy is executed correctly and without errors.
- **7.2 Performance Monitoring:** The plan for ongoing "health checks" of the live strategy, including tracking daily profit/loss and key risk metrics.
- **7.3 Failure Mode Planning:** A critical emergency plan that defines, in advance, the specific red lines (e.g., "drawdown exceeds 25%") that will trigger an immediate halt to the strategy. This prevents emotional decision-making during market downturns.

Phase 8: Documentation & Reporting

Goal: To package and communicate the research and findings effectively.

- **8.1 Research Documentation:** The internal "lab notebook" for the project. It documents every decision, rationale, and finding to ensure the work is reproducible and the methodology is transparent.
- **8.2 Academic/Professional Presentation:** The final, polished report for an external audience. It includes a concise executive summary for busy stakeholders, detailed technical results with charts and tables, and an honest assessment of the strategy's limitations and potential avenues for future research.