📘 Keystone Project Blueprint

Self-Exciting Pairs Trading with Jump-Aware Entry and Exit

1. Project Summary

This project develops a **statistical arbitrage trading strategy** that extends traditional mean-reversion pairs trading by incorporating **self-exciting jump risk dynamics** using a hybrid **Mean-Reverting Jump Diffusion (MRJD) + Hawkes process** model.

The strategy models the spread between cointegrated assets as a mean-reverting process that experiences clustered discontinuities caused by liquidity shocks, information cascades, and market stress. Instead of assuming spread deviations revert independently, the strategy dynamically adjusts trading signals based on **estimated jump intensity**, allowing the system to:

- Avoid entering trades during cascade risk
- Dynamically scale exposure based on market instability
- Improve tail-risk control and drawdown resilience

The project aims to demonstrate how structural jump clustering influences statistical arbitrage profitability and risk characteristics.

2. Project Motivation

2.1 Limitations of Traditional Statistical Arbitrage

Classical pairs trading assumes:

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t$$

Where spread deviations revert smoothly toward equilibrium.

However, empirical observations show:

• Spread deviations cluster during liquidity crises
• Large dislocations trigger additional dislocations
• Mean reversion temporarily fails during stress
• Stat-arb strategies often suffer catastrophic drawdowns during these regimes

These failures arise because traditional OU models assume:

- Independent shocks
- Continuous diffusion
- Stable reversion speed

2.2 Market Reality: Cascading Liquidity Shocks

In real markets:

- Margin calls trigger forced liquidation
- Market makers withdraw liquidity simultaneously
- ETF and derivative arbitrage breaks down
- Spread dislocations propagate across assets

These mechanisms create **self-exciting instability**, which is captured by Hawkes processes.

2.3 Research Objective

This project aims to answer:

> Can modeling clustered jump risk improve statistical arbitrage timing, profitability, and tail-risk control?

3. Model and Algorithm Purpose

The model serves three purposes:

1. Spread Dynamics Modeling

Represent equilibrium forces and shock behavior realistically.

2. Instability Forecasting

Estimate real-time cascade risk using Hawkes intensity.

3. Trading Signal Enhancement

Use jump risk estimates to adapt:

- Entry timing
- Position sizing
- Exit timing
- Risk control

4. Mathematical Framework

4.1 Spread Construction

Define spread between assets A and B:

$$X_t = \log(P_A(t)) - \beta \log(P_B(t))$$

Where:

- P_A, P_B = asset prices
- β = hedge ratio (estimated via cointegration or regression)

4.2 Hybrid Spread Model

Spread follows MRJD with Hawkes jump arrivals:

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t + Y dN_t$$

Where:

Variable	Meaning
X_t	Spread value
κ	Mean reversion speed
θ	Long-run equilibrium
σ	Diffusive volatility
W_t	Brownian motion
N_t	Jump counting process
Y	Jump magnitude

4.3 Self-Exciting Jump Process

Jump arrival intensity:

$$\lambda_t = \bar{\lambda} + \sum_{t_i < t} \alpha e^{-\beta_H (t - t_i)}$$

Where:

Parameter	Meaning
$\bar{\lambda}$	Baseline jump probability
α	Jump excitation strength
β_H	Excitation decay rate

Interpretation:

- Each spread jump increases probability of future jumps.
- Cascade risk decays gradually over time.

4.4 Trading Signal Logic

Spread Z-Score

$$Z_t = \frac{X_t - \theta}{\hat{\sigma}_X}$$

Jump-Aware Entry Condition

Enter trades when:

$$|Z_t| > Z_{entry} \quad \text{AND} \quad \lambda_t < \lambda_{threshold}$$

Dynamic Position Scaling

$$w_t \propto \frac{Z_t}{1 + c\lambda_t}$$

Meaning:

- Exposure decreases during instability regimes

5. Mathematical Intuition

The hybrid model captures two interacting forces:

Equilibrium Force

$$\kappa(\theta - X_t)$$

Represents fundamental arbitrage convergence.

Instability Force

$$Y\,dN_t$$

Represents liquidity and informational shocks.

Self-Excitation

$$\lambda_t$$

Captures contagion and clustering of shocks.

The trading strategy exploits the observation:

> Mean reversion remains valid long-term but is temporarily unreliable during clustered shock regimes.

$\downarrow$

6. Candidate Asset Pairs

Selection criteria:

- Economic linkage
- High liquidity
- Structural arbitrage relationship
- Historical cointegration

Equity Pairs

Sector Competitors

- Coca-Cola vs Pepsi
- Exxon vs Chevron
- Visa vs Mastercard

ETF vs Constituents

- SPY vs IVV
- QQQ vs XLK

Futures vs Spot

- Gold Futures vs GLD
- Oil Futures vs USO

Commodities (Highly Recommended for Differentiation)

- Henry Hub Natural Gas Futures (front vs next contract)
- Power vs Gas spreads
- Crack spreads

7. Overall Project Architecture

Data Layer

Responsibilities:

- Acquire historical prices
- Preprocess data
- Construct spreads
- Detect jumps

Model Layer

Spread Model Estimation

- Hedge ratio estimation
- OU parameter estimation
- Jump distribution estimation

Hawkes Calibration

- Jump time extraction
- Intensity parameter estimation
- Model validation

8. Project Structure (High-Level)

```
self_exciting_pairs_trading/
│
├── data_pipeline/
├── jump_detection/
├── hawkes_model/
├── mrjd_model/
├── spread_construction/
├── trading_signals/
├── backtesting/
├── risk_analytics/
├── visualization/
├── notebooks/
├── report/
```

Copy code

9. Evaluation Framework

Performance Metrics

- Sharpe Ratio
- Sortino Ratio
- Maximum Drawdown
- Calmar Ratio
- Expected Shortfall

Strategy Comparison

Compare:

1. Classical OU pairs trading
2. Volatility-scaled OU
3. Jump-aware MRJD–Hawkes strategy

Regime Analysis

Evaluate performance during:

- Crisis periods
- High volatility regimes
- Jump cluster regimes

Signal Generation Layer

- Spread deviation computation
- Jump intensity filtering
- Entry/exit logic

Backtesting Engine

- Trade execution simulation
- Transaction cost modeling
- Position management
- Performance tracking

Risk & Performance Analytics

- Drawdown analysis
- Tail risk metrics
- Regime performance comparison

Visualization Dashboard

- Spread paths
- Jump clusters
- Intensity evolution
- Trade overlays

10. Expected Contributions

This project contributes:

1. New Trading Methodology

Jump-risk-adaptive stat-arb execution.

2. Risk Insight

Quantifies how clustering affects arbitrage reliability.

3. Empirical Evidence

Demonstrates performance improvement through instability modeling.

12. Portfolio Positioning Narrative

This project demonstrates ability to:

- Apply stochastic calculus to trading systems
- Model market instability realistically
- Design alpha strategies using structural models
- Integrate statistical modeling with risk control

13. Research Questions This Project Answers

- When does statistical arbitrage fail?
- Can jump clustering forecast breakdown of mean reversion?
- How should arbitrage exposure adapt to market instability?

14. Deliverables

- Fully modular research codebase
- Interactive visualization dashboards
- Performance research report
- Reproducible calibration framework
- Strategy benchmarking study