**1. Asset Universe & Cluster Construction**

* **Select ~10–20 liquid crypto perpetual futures (BTC, ETH, SOL, ADA, AVAX, etc.) from major exchanges (Binance, Bybit, etc.).**
* **Use PCA to reduce feature dimensionality across:**
  + **Price returns**
  + **Order book imbalance**
  + **Funding rate volatility**
  + **Realized volatility**
* **Apply monthly K-Means clustering on PCA output to group coins by latent co-movements.**
* **Cluster-level goal: form baskets where members share statistically stable behaviours.**

**Filters applied per pair in each cluster:**

* **ADF test p-value < 0.05 (stationarity)**
* **Spread half-life < 12 hours (mean reversion potential)**
* **Rolling volatility (std dev) used for sizing**
* **GARCH(1,1) alpha + beta < 0.9 (volatility not too persistent)**
* **Average hourly volume > $1M USD equivalent**
* **Funding rate and volatility to avoid extreme bias**

**2. Pair Filtering & Rolling Cointegration**

* **For every cluster, test all pairwise combinations using:**
  + **Engle-Granger or Johansen cointegration tests**
  + **Rolling window (e.g. 30–60 days) to detect persistence over time**
* **Reject pairs where cointegration breaks > 30% of time window.**
* **Maintain a live whitelist of dynamically valid spreads.**

**3. Spread Construction & Signal Logic**

* **Compute log spread:**

**using OLS or Johansen's eigenvector.**

* **Compute z-score of spread using a 60-bar rolling window:**
* **Zt​=St​−μ​/σ**

**Trading signals:**

* + **Enter when |Z| > 1.5**
  + **Exit when Z ≈ 0 or max holding time exceeded (e.g. 48 bars)**
  + **Stop-loss if Z hits 2.5 or inverse volatility threshold breached**

**4. Risk Management & Sizing**

* **Position size ∝ inverse of spread's rolling volatility**

**Size∝1σspread,rolling\text{Size} \propto \frac{1}{\sigma\_{spread, rolling}}Size∝σspread,rolling​1​**

* **Optional: cap exposure by available liquidity (e.g. < 5% of avg 1h volume)**
* **Avoid trading near funding payment timestamps (e.g. 1–2 mins before/after)**
* **Suspend trading during high-volatility regime shifts (e.g. > 99th percentile spread std dev)**

**5. Execution**

* **Use limit or post-only orders to reduce taker fees and slippage.**
* **Monitor bid-ask spread and order book depth before placing trades.**
* **Cancel unfilled entries if spread mean-reverts pre-fill.**

**6. Backtesting & Evaluation**

* **Run backtests on 1-min or tick data with full costs:**
  + **Fees (maker/taker)**
  + **Slippage (based on order book depth model)**
  + **Latency assumptions**
* **Evaluate:**
  + **Sharpe & Sortino**
  + **Max drawdown**
  + **Hit rate**
  + **Trade PnL distribution**
  + **Rolling performance by volatility regime**
* **Perform walk-forward analysis and Monte Carlo stress testing**

**7. Optional Visual Dashboard**

* **Minimal dashboard (e.g. Streamlit or lightweight HTML) for:**
  + **Live signal tracking**
  + **Historical PnL & spread z-score charts**
  + **Pair whitelist management**
  + **Parameter toggling & live risk metrics**