Cointegration-Based Statistical Arbitrage on Indian Banking Stocks (HDFC & ICICI)

Author: Siddharth Sunil Bhosale
Date: 2025

---

1. Project Overview

This project implements a **statistical arbitrage (pairs trading)** strategy using the concept of **cointegration** between **HDFC Bank (HDFCBANK.NS)** and **ICICI Bank (ICICIBANK.NS)**.
The idea is that both banks operate in the same sector and tend to move together in the long term. When their prices diverge temporarily, the strategy takes advantage of **mean reversion** to generate profit.

---

2. Concept of Cointegration

Cointegration means that even if two stock prices individually follow non-stationary random walks, a **linear combination** of them can be **stationary** (mean-reverting).

Mathematically:

$$HDFC_t = \alpha + \beta \cdot ICICI_t + \varepsilon_t$$

Here:

- α= intercept
- β= hedge ratio (estimated via OLS regression)
- ε_t= residual or **spread**

If the spread (ε_t) is stationary, HDFC and ICICI are **cointegrated**.

---

3. Steps in Implementation

(a) Data Collection

- Source: Yahoo Finance
- Interval: Daily Adjusted Close Prices
- Duration: 3 years (252 trading days per year)

(b) Hedge Ratio via OLS Regression

$$HDFC_t = \alpha + \beta \cdot ICICI_t + \varepsilon_t$$

- Estimate β using Ordinary Least Squares (OLS).
- Compute the spread:

$$Spread_t = HDFC_t - \beta \cdot ICICI_t$$

(c) Stationarity Check

Perform **Augmented Dickey–Fuller (ADF) test** on the spread.

- H_0: Spread is non-stationary
- H_1: Spread is stationary
- If p-value < 0.05 → reject H_0 → spread is stationary → cointegration confirmed.

(d) Z-score Calculation

To standardize the spread:

$$Z_t = \frac{Spread_t - \mu_{Spread,60}}{\sigma_{Spread,60}}$$

Where:

- $\mu_{Spread,60}$ = 60-day rolling mean
- $\sigma_{Spread,60}$ = 60-day rolling standard deviation

---

4. Trading Logic

Condition	Action	Explanation
$Z > +2$	Short Spread	HDFC overpriced vs ICICI
$Z < -2$	Long Spread	HDFC underpriced vs ICICI
(	Z	< 0.5)

- **Long spread:** Buy HDFC, Sell β·ICICI
- **Short spread:** Sell HDFC, Buy β·ICICI

---

5. Backtesting Logic

1. Compute daily change in spread:

$$\Delta Spread_t = Spread_t - Spread_{t-1}$$

2. Compute strategy returns:

$$R_t = Position_{t-1} \cdot \Delta Spread_t$$

3. Normalize returns by volatility to make them comparable.
4. Subtract small transaction costs (4 basis points per trade).
5. Evaluate performance metrics:

- **Sharpe ratio:**

$$Sharpe = \frac{E[R]}{\sigma_R}\sqrt{252}$$

- **Max Drawdown:** largest drop from equity peak.
- **Trade counts:** how many long and short positions were executed.

---

6. Results Summary

Metric	Value
Hedge ratio (β)	≈ 1.12
ADF p-value	0.017 (<0.05 → stationary)
Sharpe Ratio	2.28
Max Drawdown	–2.5
Trades Executed	16 (9 long, 7 short)

Inference: The spread between HDFC and ICICI is **stationary and mean-reverting**, validating the cointegration hypothesis.

---

7. Challenges & Solutions

Challenge	Explanation	Solution
Overnight Gap Risk	Sudden price jumps next morning can cause losses.	Developed **intraday-only version** to close all trades by market close.
Low Volatility Periods	Spread hardly moves, causing low returns.	Applied **volatility filter** — trade only between 10th–90th percentile of spread volatility.
High Volatility / Earnings Events	Sudden spikes distort z-score.	Skip trades on abnormal days (>3σ spread moves).
Non-stationary periods	Market regimes shift, breaking equilibrium.	Recalculate **hedge ratio β** periodically (rolling regression).

---

8. Conclusion

The HDFC–ICICI pair shows strong cointegration, and the mean-reverting spread can be exploited for short-term arbitrage.
The strategy achieved a **Sharpe ratio of 2.28** with **limited drawdowns**.
Risk controls like volatility filters and intraday execution further improve robustness.