Analyzing Indian Stock Markets through Correlations: Comparative Insights

Ashish Shrivastav[1], Stuti Patel[2], Vihaan Kapopara[3], Archita Chandwani[4], Mann Ahalpara[5], Ashlesha Bhise[6]

[1,2,3,4,5] Information and Communication Technology, Adani University, Gujarat, India
[6] Faculty of Sciences, Adani University, Gujarat, India
AshishShrivastav.ict23@adaniuni.ac.in

Abstract: Traditional methods for analyzing emerging markets like India often fall short in capturing the full picture of systemic risk because they rely on single-metric, linear correlation models. To address this gap, this study introduces a multi-metric comparative framework built on concepts from Financial Engineering and Data Science to examine how ten major sectors of the Indian stock market move together over time. The approach integrates four different correlation measures—Pearson, Spearman, Kendall's Tau, and time-lagged cross-correlation—to capture not just linear connections but also monotonic and time-dependent relationships. These correlations are then combined with Minimum Spanning Trees (MSTs) and clustering algorithms to map out the network of sector relationships and track how it evolves with market conditions. The results show a significant average divergence (AvgDiff ≈ 0.40) among the different correlation methods, highlighting that relying on a single measure can give an incomplete view of market dynamics. The analysis effectively grouped the market into distinct clusters: the Automobile sector emerged as a tightly connected network (Pearson $\rho \approx 0.96$), whereas the pharmaceutical sector appeared more dispersed, indicating greater potential for diversification. Furthermore, the time-lagged analysis revealed lead-lag effects, such as a 75-day delay between specific stock movements. Overall, this study provides a comprehensive and data-driven framework for assessing systemic risk and optimizing portfolios in volatile emerging markets like India, offering deeper insights into how different sectors interact and influence each other over time

Keywords: Indian equity systems, inter-market relationships, diversification insights, risk assessment, network modeling

R. Mehta et al. (eds.), *Proceedings of the International Conference on Infrastructure Development and Sustainability (ICIDS 2025)*, Atlantis Highlights in Sustainable Development 9,
https://doi.org/10.2991/978-94-6239-685-2_13

1 Introduction

The stock market plays a crucial role in driving economic growth, serving as both a mirror of a nation's economic ambitions and the uncertainties of its financial environment. In India, this journey began with the establishment of the Bombay Stock Exchange in 1875, which has since developed into one of the most dynamic markets in the world. Today, the Indian stock market functions as a key financial hub, embodying the country's economic objectives, structural diversity, and growing global integration. Its intricate nature makes it especially compelling: changes in one part of the market can create ripples across others, influencing profits and losses, and this interdependence is crucial for overall financial stability.

Traditional approaches to market analysis often simplify these interconnections. Fundamental analysis—through tools like Discounted Cash Flow (DCF), Price-to-Earnings (P/E) ratios, Dividend Discount Models, and macroeconomic indicators—focuses on estimating a stock's intrinsic value but fails to capture how different stocks interact. Technical analysis, with tools such as Moving Averages, Relative Strength Index (RSI), Bollinger Bands, and MACD, helps identify trading opportunities based on historical patterns but does not explicitly represent cross-stock or cross-sector relationships. While both methods provide meaningful insights, they mainly emphasize individual indicators rather than the wider network of market interconnections. The study of interactions between various stocks and their influence on one other remains challenging and largely rely on various methods of correlation as discussed in existing literature.

In financial research, correlation analysis remains important. It reveals the relationships and trends in stock markets [1], [2]. Although Pearson correlation is the most widely used, numerous studies demonstrate that rank-based and non-linear correlations, such as Spearman and Kendall, can uncover associations that linear models fail to detect [3], [4]. Correlation serves as a fundamental tool in financial risk assessment [5]. Its applications extend to partial correlation networks in emerging markets [6] and to examining connection between currencies and commodities [7]. Recent research has also associated investor sentiment with stock price fluctuations, further broadening the use of these techniques [8], [9]. It also applies clustering techniques to find hidden communities [10] and detrended cross-correlation [11], [12] to explore temporal dynamics.

Despite these advances, most traditional financial analyses remain heavily dependent on the Pearson coefficient, which captures only linear relationships. This single-metric dependence fails to account for non-linear, rank-order, and time-lagged dependencies that are prominent in high-noise, information-asymmetric EMs. Consequently, systemic risk is often underestimated, and diversification strategies become less effective. A more comprehensive analytical approach is therefore required to accurately represent market co-movements and risk dynamics.

To overcome these limitations, this study proposes a novel multi-metric comparative framework. This study makes the following contributions to the literature on financial dependence and systemic risk analysis:

1. Proposes a multi-metric dependence framework that jointly analyzes linear, rank-based, and temporal correlations within a unified structure.

2. Introduces a quantified deviation metric (AvgDiff) to measure instability and divergence across dependence measures, highlighting limitations of single-metric analysis.

3. Demonstrates empirically that sectoral integration in the Indian market is dimension-dependent, varying significantly across linear, monotonic, and temporal measures.

4. Identifies long-horizon lead–lag effects (up to 75 trading days) between cyclical and defensive sectors, revealing macro-financial signaling mechanisms not observable through static correlations.

5. Provides actionable implications for systemic risk assessment, diversification strategy, and time-adaptive portfolio design in emerging markets.

Focusing on ten key sectors of the Indian stock market—including Banking, Automobiles, and Pharmaceuticals—the framework integrates four distinct correlation measures: Pearson, Spearman, Kendall's Tau, and time-lagged cross-correlation. These are further analyzed using Minimum Spanning Trees (MSTs) and dynamic clustering algorithms to uncover the evolving topological structure of inter-sectoral relationships.

The remainder of this paper is structured as follows. Section 2 reviews relevant literature on financial network analysis and correlation techniques. Section 3 outlines the data and the multi-metric methodology employed. Section 4 presents and discusses empirical results; and Section 5 concludes the study and suggests avenues for future research.

2 Methodology

Our work uses four different techniques to evaluate how stock returns are connected. Each technique showed a unique way in which stock prices relate to each other, and together they showed different kinds of relationships—like steady, straight-line, rank-based, or time-related links. The methodology integrates four complementary statistical approaches—**Pearson, Spearman, Kendall's Tau,** and **Cross-Correlation**—to capture linear, non-linear, and temporal associations among sectoral stock returns. It offers a detailed overview of the key concepts, theories, and mathematical formulas, ensuring that the analysis is conducted consistently and understood the same way every time.

All the four correlation measures—Pearson, Spearman, Kendall's Tau, and Cross-Correlation—range from -1 to +1. A value of +1 indicates a perfect positive relationship, meaning if one stock's return increases, the other also increases, and if one decreases, the other also decreases. A value of -1 indicates a perfect negative relationship, meaning if one stock's return increases, the other decreases, and vice versa. A value of 0 indicates no relationship between the stock returns.

Common Setup.

The initial action was to compile the data and transform the prices into a format that could be used for detailed analysis. Initially, the closing prices of ten large businesses from ten sectors were collected and organized by the day of trading to ensure data consistency. Raw prices of daily data were made more stable, making it easier to check the correlation between different stocks. The computational data can be evaluated as either a simple return:

Since the dataset is already cleaned and prepared, no further cleaning is required. Outliers show real market changes, and rank-based measures like Spearman and Kendall are still reliable even with them.

Let $\{X_i\}^n$ and $\{Y_i\}^n$ indicate the return series of two stocks X and Y over exchange days. These return series are then used to derive the correlation matrices.

2.1 Pearson Correlation — Linear Co-Movement

The Pearson correlation is the most common way to measure how closely two stocks move together in a straight line. It checks if the changes in one stock's returns are in the same proportion as the changes in another. To calculate it, we divided the covariance of the two sets of returns by the product of their standard deviations.

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \, . \, \sigma_Y}$$

where,

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \underline{X})(Y_i - \underline{Y})$$

The Pearson correlation amount ranges from –1 to +1. A positive value indicates that two stocks move in the same direction in proportion, while a negative value indicates they shift in opposite directions. Values close to zero suggest there is no clear linear relationship between the stocks. In this study, Pearson correlation served as the foundation for understanding proportionate co-movements among stocks, enabling the identification of simple stock connections, such as IT companies moving together following news affecting the entire sector. The advantages and the disadvantages of the Pearson Correlation are given in Table 1.

Table 1. Advantages and disadvantages of Pearson Correlation

Advantages	Disadvantages
- Simple to compute and widely used in finance.	- Captures only linear relationships, missing nonlinear dependencies.
- Measures the strength and direction of linear relationships.	- Sensitive to outliers, which can distort results in volatile markets.
- Useful when stock returns move proportionally (up or down) together.	- Assumes normally distributed data, which may not be held in stock returns.
- Provides easily interpretable values between –1 and +1.	

2.2 Spearman's Rank Correlation — Monotonic Co-Movement

Spearman's rank correlation is a way to look at how two stocks move together over time, not just in a straight line. It checks if the two stocks tend to go up or down together, even if they don't move exactly the same way. This approach was not greatly affected by extreme values because ranking lowers the effect of outliers. In financial markets, this was beneficial as stock prices can move in the same direction even if their price changes vary significantly. In our analysis, Spearman's method helped detect similar trends across various sectors. For instance, it showed that if HDFC Bank usually increases on most days, ICICI Bank often follows the same trend, regardless of the actual price movement.

$$\rho_s = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$$

where $d_i = R(X_i) - R(Y_i)$ is the difference between the ranks of X and Y.

This approach was less affected by extreme values because ranking minimizes the impact of outliers. In financial markets, this is especially significant, as stocks can move in the same direction even if their price changes differ greatly. In our analysis, Spearman's method helped detect consistent directional trends across different sectors, for example, indicating that if HDFC Bank tends to go up on most days, ICICI Bank often follows a similar pattern, regardless of the specific percentage changes. The advantages and the disadvantages of the Spearman's Rank Correlation are given in Table 2

Table 2. Advantages and disadvantages of Spearman's Rank Correlation

Advantages	Disadvantages
- Non-parametric; does not assume normality of data.	- Less precise than Pearson when relationships are strongly linear.
- Captures monotonic relationships (stocks moving in the same direction, even if not at the same rate).	- May lose information about actual magnitudes of movement (focuses on ranks).
- Less affected by extreme outliers compared to Pearson.	- Can be less stable with small datasets.
- Useful for ranking-based comparisons across stocks.	

2.3 Kendall's Tau — Concordance/Discordance of Order

Kendall's Tau correlation is additional rank-based method but depends on counting concordant and discordant data pair. For two time points i and j, the pair is concordant if both the stocks display the same return ordering, and discordant if the order is opposite. The formula is:

$$\tau = \frac{(C - D)}{\frac{1}{2}n(n-1)}$$

where *C* is the number of concordant pairs and *D* is the number of discordant pairs. A tie-adjusted version, Kendall's τ_b, is often used in practice.

Kendall's Tau is more reluctant to noisy data compared to Spearman because it focuses on the agreement in the order of rankings rather than the actual size of the differences. It essentially measures the probability that two stocks are moving in the same direction minus the probability they are moving in contrary directions. In our study, this helped highlight the long-term consistency in the relationships between stocks, such as whether two pharmaceutical companies like Sun Pharma and Cipla tend to rise and fall together, even during periods of market volatility. The advantages and the disadvantages of the Kendall's Tau are given in Table 3

Table 3 Advantages and disadvantages of Kendall's Tau

Advantages	Disadvantages
- Robust, non-parametric, and highly resistant to noise and outliers.	- Computationally more intensive than Pearson or Spearman.
- Measures strength of association based on concordant/discordant pairs.	- Produces smaller correlation coefficients (making interpretation less intuitive).

- Provides more reliable estimates for smaller sample sizes. - Good for identifying long-term directional associations.	- Less sensitive to subtle short-term changes in stock returns.

2.4 Cross-Correlation — Lead–Lag Dependencies

While Pearson, Spearman, and Kendall focus on measuring relationships at a single point in time, cross-correlation adds a time element by examining how the return of one stock today might relate to the return of another stock tomorrow, or the other way around. It is defined as:

$$\rho XY(k) = \frac{\sum_t (X_t - \underline{X})(Y_{t+k} - \underline{Y})}{\sqrt{\sum (X_t - \underline{X})^2 \sum (Y_{t+k} - \underline{Y})^2}}$$

where k represents the lag. Positive k indicates that X leads Y, while negative k means Y leads X. By calculating this over numerous lags (e.g., –5 to +5 days), we can detect *leader–follower dynamics.*

This method was helpful for finding out how time affects balance between stocks, like when one stock goes down first and another goes up a little later, which helps keep the whole industry stable. In our study, cross-correlation showed situations where changes in one company or sector happen a bit later in others. The advantages and the disadvantages of the Cross-Correlation are given in Table 4 .

Table 4 Advantages and disadvantages of Cross Correlation

Advantages	Disadvantages
- Captures time-shifted relationships across sectors, identifying which stocks lead and which follow.	- Requires longer time series for reliable estimation.
- Useful for detecting balancing effects and sectoral dependencies over time.	- More sensitive to noise in high-frequency financial data.
- Adds temporal dimension to correlation analysis, often missed in static methods.	- Interpretation can be complex, especially when multiple lags show significant correlations.
- Valuable for understanding systemic risk and sector rotations.	- Risk of spurious correlations if not properly filtered.

Outcome

These four methods provided various ways to understand how stocks are interconnected with each other. Pearson's method demonstrated how stocks move in correlation to one another. Spearman and Kendall looked at how the rankings of stocks are consistently linked, even when the relationship wasn't straight or when there were odd data points. Cross-correlation helped determine which stocks usually come before or after others. Combined, these methods gave an extensive view of how different stocks within Indian sectors are related.

Econometric Interpretation of Dependence Measures

The four correlation measures employed in this study are grounded in distinct econometric interpretations of dependence:

1. **Pearson correlation** captures linear dependence and is directly aligned with classical mean–variance portfolio theory, where proportional co-movement determines diversification efficiency.
2. **Spearman's rank correlation** measures monotonic dependence and is robust to non-normality and outliers, reflecting dependence structures analogous to copula-based relationships without imposing parametric assumptions.
3. **Kendall's Tau** quantifies concordance probability and provides a more conservative estimate of rank dependence, particularly suitable for noisy and finite samples.
4. **Cross-correlation** introduces a temporal dimension, enabling detection of lead–lag relationships and information transmission across assets, which are central to understanding systemic propagation mechanisms.

Joint analysis of these measures allows the decomposition of market dependence into linear, monotonic, and temporal components, offering a more complete econometric characterization of inter-stock relationships than any single metric in isolation.

3 Quantitative Analysis

3.1 Anchor Stock Selection and Analytical Rationale

To comprehensively evaluate the robustness of inter-stock relationships, we computed the absolute deviations across different correlation measures: Pearson (ρ_P), Spearman's rank correlation (ρ_S), Kendall's tau (τ_k), and Cross-Correlation (ρ_{cc}). The Average Absolute Deviation from the Pearson coefficient was quantified using the AvgDiff metric, defined as:

$$AvgDiff = \frac{|\rho_P - \rho_S| + |\rho_P - \tau_K| + |\rho_P - \rho_{CC}|}{3}$$

The results indicate substantial and frequent deviations, with AvgDiff values observed to range from approximately 0.35 to 0.50 across the top deviating company pairs. This narrow, high-magnitude range suggests that the linear, monotonic, and lag-based relationships captured by these methods often differ significantly. The highest overall average deviation is observed in the Real Estate–Pharma sector pair (PHOENIXLTD.NS – GLENMARK.NS), which shows an AvgDiff of 0.453.

This analysis highlights a critical structural difference between our anchor stocks:

1. **GLENMARK.NS (Pharma):** Pairs involving Pharma stocks frequently appear among the top deviation scores. This suggests that the Pharma sector, characterized by its low linear correlation, exhibits stronger monotonic or lagged co-movements with other industries that are not captured linearly.

2. **BAJAJ-AUTO.NS (Auto):** In contrast to the low intra-sector AvgDiff observed within the tightly coupled Auto segment, BAJAJ-AUTO.NS appears high in the deviation list primarily in cross-sector combinations (e.g., Auto–Pharma AvgDiff ≈0.44; Auto–Telecom AvgDiff ≈0.38). This distinction underscores that while Auto is internally synchronous, its relationships with heterogeneous sectors rely heavily on non-linear or lag-based dynamics, causing Pearson to be a poor estimator of total dependence.

The Necessity of a Multi-Metric Approach

Our current analysis of stock correlation has thus far focused predominantly on the Pearson correlation method, allowing for deep dives into diversification and momentum gain strategies based on linear relationships. However, the significant AvgDiff results confirm that relying on a single correlation measure cannot fully capture the complex sectoral and cross-sector dynamics observed in the market. Moving forward, the incorporation of linear (Pearson), non-linear/rank-based (Spearman, Kendall's Tau), and lag-sensitive (Cross-Correlation) methods is crucial to provide a more comprehensive and solid grasp of inter-stock relationships.

3.2 Graphical Overview and Interpretation of Inter-Stock Relationships

This section provides a graphical analysis of inter-stock relationships within the Banking sector, with a focus on ICICIBANK.NS as a representative stock. The objective is to illustrate the methodology and interpretative framework in a manner accessible to both technical and non-technical audiences. The analysis demonstrates how these visualizations offer actionable insights for investment strategies, such as momentum trading and diversification, which will be extended to the comparative study of GLENMARK.NS and BAJAJ-AUTO.NS in subsequent sections.

Correlation Heatmap: Measuring Linear Co-Movement

The Correlation Matrix displays the Pearson correlation coefficients (ρ) between all pairs of stocks in the Banking sector, ranging from -1 to 1. This tool measures the degree of linear co-movement in stock prices.

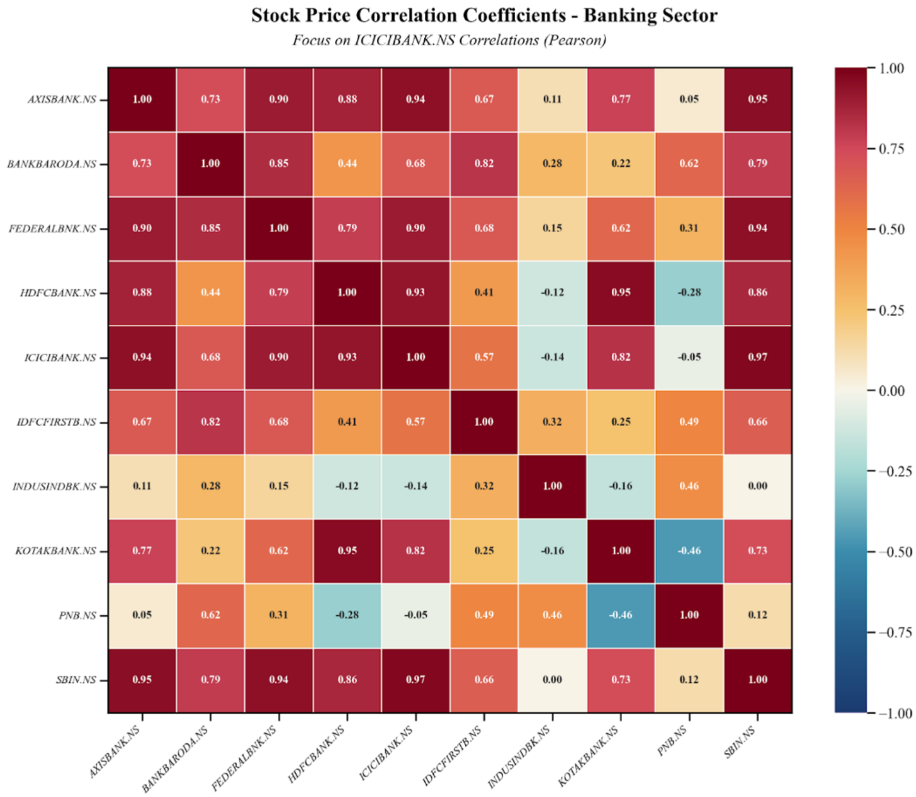

Fig. 1. Stock Price Correlation Coefficients - Banking Sector (Pearson)

Interpretation of Figure 1: The matrix is constructed by calculating the Pearson correlation coefficients from daily price changes over the analysis period. A value near 1 (represented by darker red hues) indicates strong positive linear alignment, while a value close to 0 (represented by lighter shades) suggests independence. For instance, ICICIBANK.NS shows a high correlation of (ρ ≈0.97) with SBIN.NS, reflecting nearly identical linear price trends, but a negligible correlation of (ρ ≈ -0.05) with PNB.NS, indicating minimal linear interdependence.

High correlations support momentum strategies where stocks are expected to move synchronously, while low correlations are essential for diversification, as they reduce exposure to common systemic risks.

Stock Correlation Network: Visualizing Structural Links

The Stock Correlation Network visualizes the linear relationships between ICICIBANK.NS and its sector peers. This diagram simplifies the matrix data by plotting stocks as nodes and drawing links (lines) based on the correlation strength.

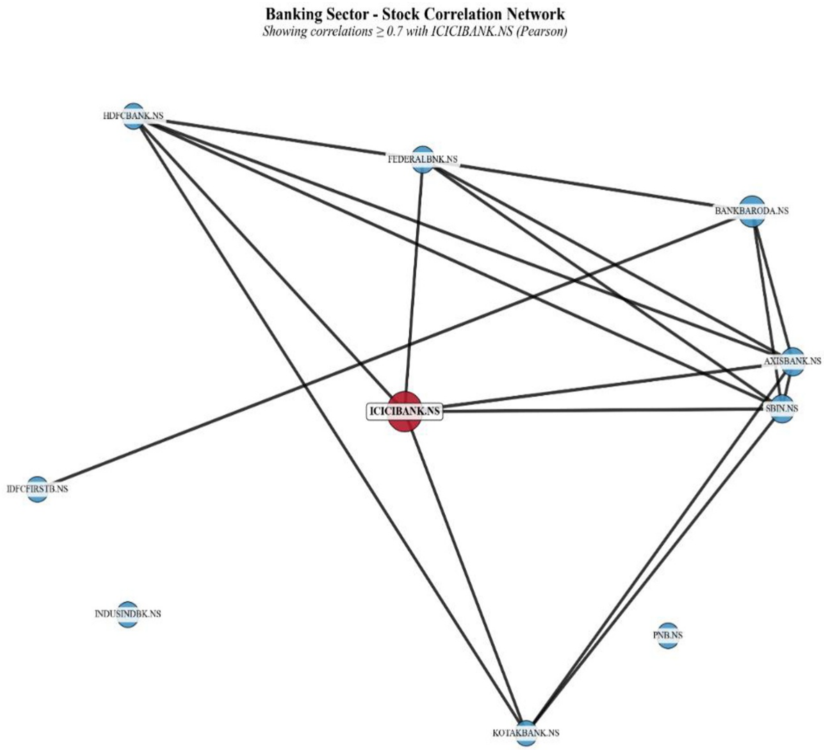

Fig.2. Banking Sector - Stock Correlation Network (Showing correlations ≥0.7 with ICICIBANK.NS (Pearson))

Interpretation of Figure 2: Links are drawn based on correlation data from the matrix, with the line color and thickness reflecting correlation strength. The diagram primarily includes only those links with a Pearson correlation of ρ≥0.7 (High Correlation). ICICIBANK.NS acts as the central node, connected to highly correlated stocks like SBIN.NS, HDFC Bank (HDFCBANK.NS), and Kotak Mahindra Bank (KOTAKBANK.NS) by strong links. The network clearly highlights the strong co-movement among the major private and public banks.

Strong links (e.g., ρ≈0.97 with SBIN.NS) are ideal for momentum trading due to high co-movement certainty. Conversely, the isolation of certain nodes, such as INDUSINDBK.NS and PNB.NS (ρ≈-0.05 with ICICIBANK.NS from the bar chart), demonstrates where risk reduction through diversification is most effective.

Intra-Sector Correlations Bar Chart: Categorizing Strategy

The Intra-Sector Correlations Bar Chart (Figure 3: Bar Chart) plots the Pearson correlation coefficient for each Banking stock relative to the anchor stock, ICICIBANK.NS, clearly categorizing the relationships based on defined strength thresholds. This visualization provides the clearest strategic insight for identifying diversification and momentum opportunities.

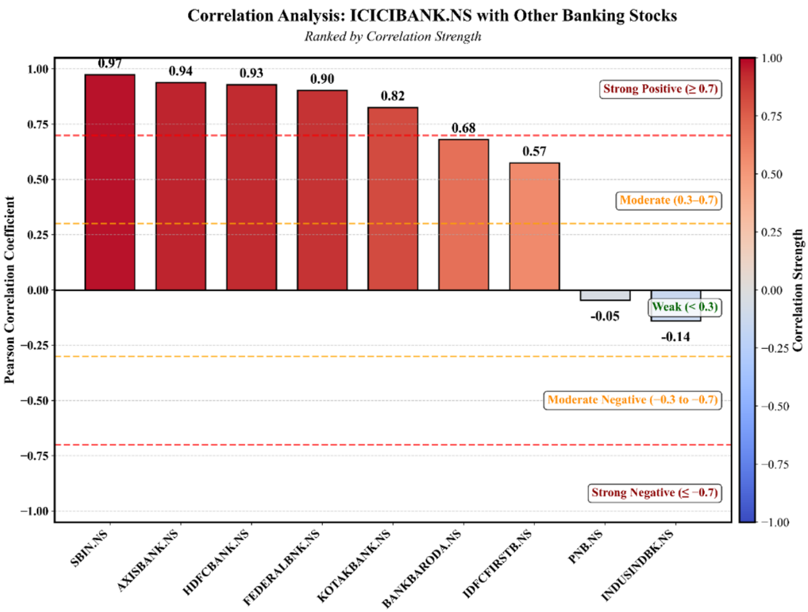

Fig.3. Correlation Analysis: ICICIBANK.NS with Other Banking Stocks (Ranked by Correlation Strength)

Interpretation of Figure 3: Correlation values are plotted as bars, with horizontal lines indicating strength thresholds: Strong Positive (ρ≥0.7), Moderate Positive (ρ∈[0.3, 0.7)), Weak (Negative or near zero), Moderate Negative (ρ∈(-0.7, -0.3]), and Strong Negative ρ≤-0.7.

1. Strong Alignment (Momentum): Stocks such as SBIN.NS (ρ≈0.97), AXISBANK.NS (ρ≈0.94), and HDFCBANK.NS (ρ≈0.93) exceed the ρ≥0.7 threshold.

2. Moderate Alignment: BANKBARODA.NS (ρ≈0.68) and IDFCFIRSTB.NS (ρ≈0.57) fall into the moderate positive range.

3. Weak/Inverse Alignment (Diversification): PNB.NS (ρ ≈ -0.05) and INDUSINDBK.NS (ρ ≈ -0.14) are classified as having weak/inverse relationships, as they fall below the moderate threshold.

The high-correlation stocks are excellent candidates for momentum strategies, as they are expected to move synchronously with ICICIBANK.NS. Conversely, the stocks classified as Weak or showing Negative correlation (PNB.NS and INDUSINDBK.NS) are essential for diversification, as their independent or inverse price movements reduce overall portfolio exposure to common systemic risks within the Banking sector.

Time Series Plot: Price Movement and Performance Disparity

The Time Series Plot shows the historical price movement of selected stocks over the analysis period. Crucially, the y-axis represents the actual Closing Price (INR), offering a concrete view of absolute price path and magnitude.

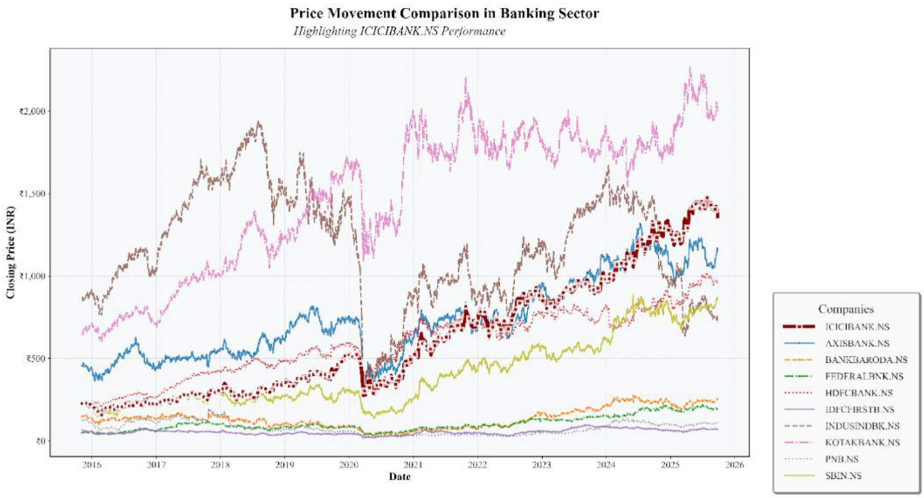

Fig. 4. Price Movement Comparison in Banking Sector

Interpretation of Figure 4: Parallel lines indicate high co-movement in price changes, visually validating a high correlation coefficient. Divergent or crossing lines indicate poor co-movement. Note that the scale is heavily influenced by high-value stocks (like HDFCFIRSTB.NS and KOTAKBANK.NS), which can visually compress the movement of lower-priced stocks (like PNB.NS).

This chart highlights that while correlation is useful for assessing risk synchronicity, it does not guarantee similar returns. The high correlation between ICICIBANK.NS and SBIN.NS implies similar risk exposure, yet their absolute price paths show distinct growth trajectories. It demonstrates that looking beyond simple correlation numbers is essential to understanding a stock's full potential for risk mitigation or momentum capture, especially when comparing stocks with large differences in price magnitude.

3.3 Pearson's Correlation – Linear Co-Movement

The Pearson correlation matrices quantify the linear co-movement within and across sectors. This analysis captures the extent to which the returns of two stocks move proportionally to one another. The results for the Auto and Pharma sectors reveal distinct patterns in their dependence structures, similar to the Spearman analysis but specifically focused on linearity.

Auto Sector Analysis (Centred on BAJAJ-AUTO.NS):

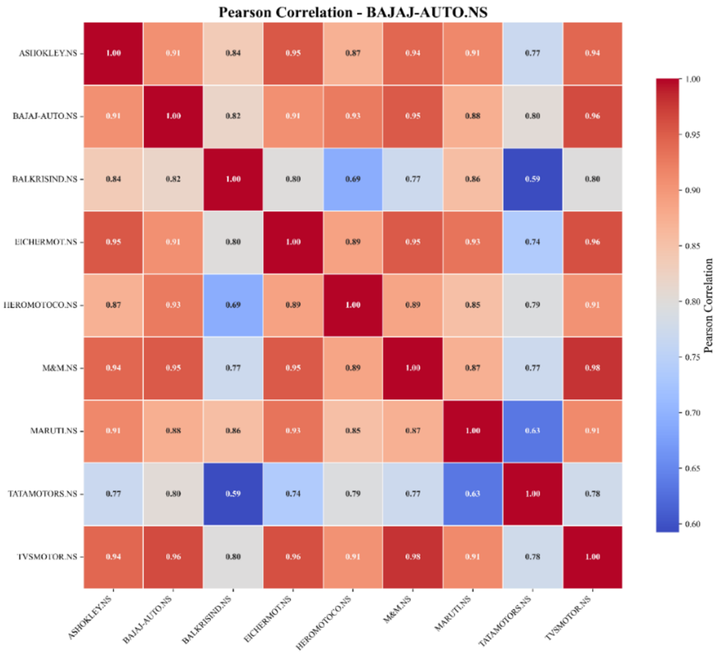

Fig. 5. Pearson Correlation - BAJAJ-AUTO.NS

Interpretation of Figure 5: The Pearson correlation matrix for the Auto sector demonstrates a highly integrated segment with uniformly high linear correlations. The values are generally elevated, indicating that the magnitude of stock returns across the sector moves in close, linear lockstep.

1. BAJAJ-AUTO.NS exhibits strong linear relationships, particularly with TVSMOTOR (ρ≈0.96), M&M (ρ≈0.95), and EICHERMOT (ρ≈0.91).

2. Several inter-auto pairs show near-perfect linear co-movement, such as TVSMOTOR–M&M (ρ≈0.98), TVSMOTOR–ASHOKLEY (ρ≈0.94), and EICHERMOT–ASHOKLEY (ρ≈0.95). These high coefficients underscore the tight linear proportionality of returns across the Auto sector, suggesting a near-uniform response to shared market and macroeconomic factors.

3. The lowest correlations are generally observed with TATAMOTORS, but even these values remain moderate to high (e.g., BAJAJ−AUTO–TATAMOTORS ρ≈0.80; BALKRISIND–TATAMOTORS ρ≈0.59), reinforcing the overall highly coupled linear structure.

Pharma Sector Analysis (Centred on GLENMARK.NS):

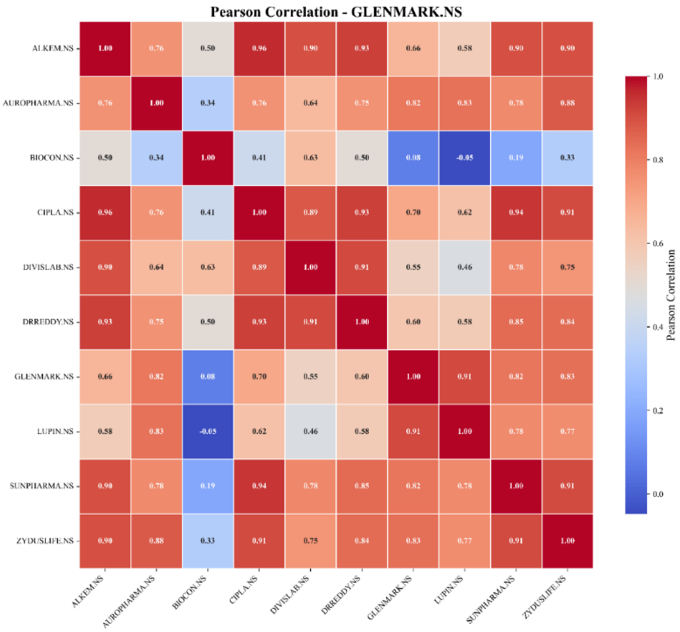

Fig. 6. Pearson Correlation - GLENMARK.NS

Interpretation of Figure 6: In stark contrast, the Pharma sector (heatmap centred on GLENMARK.NS) is significantly more heterogeneous and less linearly correlated. While strong clusters exist, the overall structure is fragmented.

1. GLENMARK.NS shows its strongest linear relationship with LUPIN (ρ≈0.91) and high correlations with AUROPHARMA (ρ≈0.82) and ZYDUSLIFE (ρ≈0.83).

2. Crucially, the relationship between GLENMARK.NS and BIOCON.NS is very weak (ρ≈0.08), indicating almost no linear relationship between their stock returns.

3. The heatmap clearly shows the coexistence of strong linear sub-clusters (e.g., ALKEM–CIPLA ρ≈0.96; CIPLA–ZYDUSLIFE ρ≈0.91) alongside weak or scattered correlations, notably involving BIOCON.NS (e.g., BIOCON–LUPIN ρ≈-0.05; BIOCON–SUNPHARMA ρ≈0.19). This structural variation implies that stock returns in the Pharma sector are driven by heterogeneous, company-specific factors that do not lead to a broad, linear co-movement.

Cross-Sector Implications

Pearson correlation coefficients for Auto sector stock pairs take values primarily in the interval [0.85, 0.96], with the maximum coefficient observed for the BAJAJ-AUTO.NS–TVSMOTOR.NS pair (ρ = 0.96). Pearson correlation coefficients within the Pharma sector span a wider interval, ranging from ρ ≈ 0.08 for the GLENMARK.NS–BIOCON.NS pair to values below ρ = 0.70 for the majority of intra-sector pairs. The numerical separation between the two sectors is therefore characterized by a lower bound of approximately 0.85 for Auto sector correlations versus values approaching 0.08 in the Pharma sector, indicating a substantially broader dispersion of linear dependence coefficients in the latter.

3.4 Spearman Rank Correlation – Monotonic Co-Movement

Spearman rank correlation coefficients for Auto sector stock pairs predominantly exceed 0.80, indicating strong alignment in return rankings across the sector. In contrast, Spearman coefficients within the Pharma sector span a wider range, from moderate positive values around 0.55 (e.g., GLENMARK.NS–LUPIN.NS) to weak or negative values near −0.12 (e.g., GLENMARK.NS–BIOCON.NS). Cross-sector Spearman coefficients are consistently lower than intra-sector Auto values and generally remain below 0.40, indicating reduced monotonic association between Auto and Pharma stocks relative to intra-sector relationships.

Auto Sector Analysis (Centred on BAJAJ-AUTO.NS)

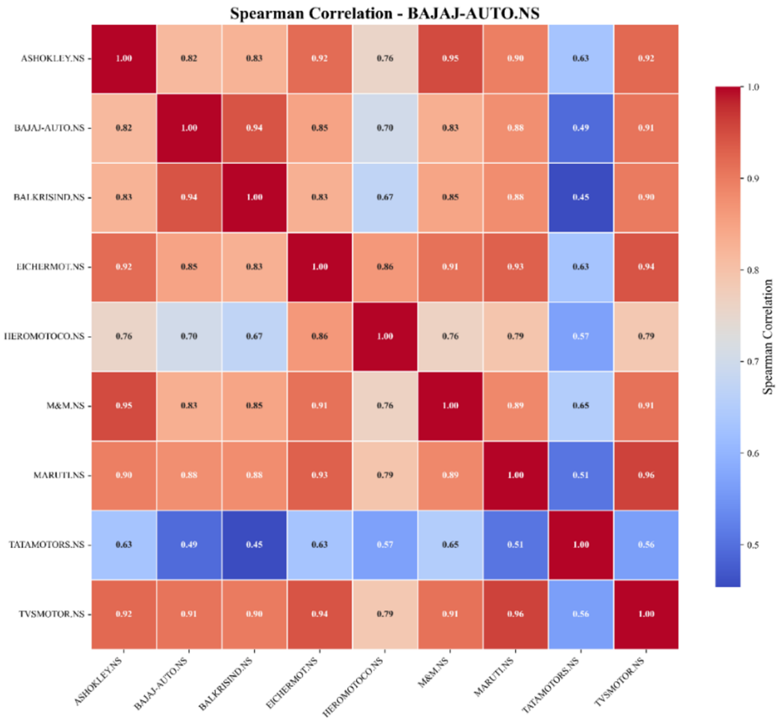

Fig. 7. Spearman Correlation - BAJAJ-AUTO.NS

Interpretation of Figure 7: The Spearman correlation matrix for the Auto sector demonstrates a highly synchronized and tightly integrated rank-based structure. The coefficients are uniformly high, indicating that changes in stock return rankings tend to move in a strongly monotonic and consistent manner across the sector.

1. BAJAJ-AUTO.NS exhibits strong monotonic associations with key peers such as BALKRISIND ($\rho_s \approx 0.93$), TVSMOTOR ($\rho_s \approx 0.89$), MARUTI ($\rho_s \approx 0.85$), and M&M ($\rho_s \approx 0.83$).

2. Several inter-auto pairs exhibit near-perfect rank synchronization, including TVSMOTOR–MARUTI ($\rho_s \approx 0.96$ and M&M–ASHOKLEY ($\rho_s \approx 0.96$), underscoring the cohesive monotonic alignment of stock return movements within the Auto sector. These high coefficients indicate that even when returns do not move in strict linear proportion, their orderings remain highly consistent in direction across market fluctuations.

Pharma Sector Analysis (Centred on GLENMARK.NS):

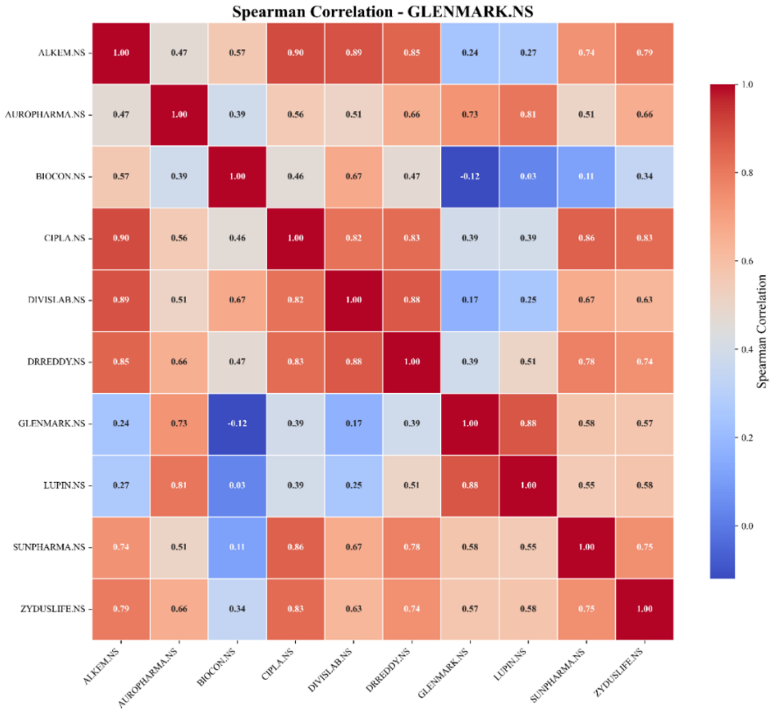

Fig. 8. Spearman Correlation - GLENMARK.NS

Interpretation of Figure 8: In contrast, the Pharma sector (heatmap centred on GLENMARK.NS) presents a more heterogeneous and fragmented rank-based dependence structure. While certain sub-clusters exhibit strong monotonic relationships, the overall pattern lacks the uniform rank synchronization observed in the Auto sector.

1. GLENMARK.NS shows a pronounced positive rank correlation with LUPIN ($\rho_s \approx 0.88$) and moderate relationships with AUROPHARMA ($\rho_s \approx 0.73$), SUNPHARMA ($\rho_s \approx 0.58$), and ZYDUSLIFE ($\rho_s \approx 0.57$). However, its relationship with BIOCON turns slightly negative ($\rho_s \approx -0.12$), suggesting a divergence in rank behaviour.

2. The heatmap further highlights well-defined sub-clusters—such as CIPLA–ALKEM ($\rho_s \approx 0.90$) and DRREDDY–DIVISLAB ($\rho_s \approx 0.88$)—alongside weak or even opposite ranking tendencies involving BIOCON and other constituents. This diversity indicates that monotonic relationships in Pharma are highly localized, driven by company-specific dynamics rather than sector-wide rank alignment.

Cross-Sector Implications

For BAJAJ-AUTO.NS, Spearman rank correlation coefficients with Auto sector peers predominantly exceed 0.80, with multiple pairwise values above 0.85, indicating consistent alignment in return rankings across the sector. For GLENMARK.NS, Spearman coefficients with Pharma sector peers exhibit a wider numerical spread, ranging from approximately 0.55 (GLENMARK.NS–LUPIN.NS) to values near −0.12 (GLENMARK.NS–BIOCON.NS), with several pairwise coefficients falling below 0.40.

Across sectors, Spearman correlation coefficients for the BAJAJ-AUTO.NS–GLENMARK.NS pair and other Auto–Pharma combinations remain below 0.40, which is substantially lower than the corresponding intra-sector Spearman coefficients observed for BAJAJ-AUTO.NS. The numerical difference between intra-sector and cross-sector rank correlations exceeds 0.40 for multiple stock pairs.

Comparison with Pearson correlation values shows that, while intra-sector Pearson coefficients in the Auto sector exceed 0.85, corresponding Spearman coefficients remain in a similar range, whereas in the Pharma sector Pearson coefficients range from approximately 0.08 to 0.70, while Spearman coefficients range from approximately −0.12 to 0.55. This numerical divergence indicates that rank-based dependence values differ substantially from linear dependence values for cross-sector and Pharma-sector stock pairs.

3.5 Kendall's Tau – Concordance and Robustness

The Kendall rank correlation matrices reveal pronounced contrasts between sectors, though the absolute coefficients are generally lower than their Spearman counterparts, consistent with Kendall's stricter concordance-based formulation. On average, Kendall values are lower by approximately 0.1–0.2 relative to Spearman for the same stock pairs, reflecting the more conservative nature of Kendall's τ.

Auto Sector Analysis (Centred on BAJAJ-AUTO.NS)

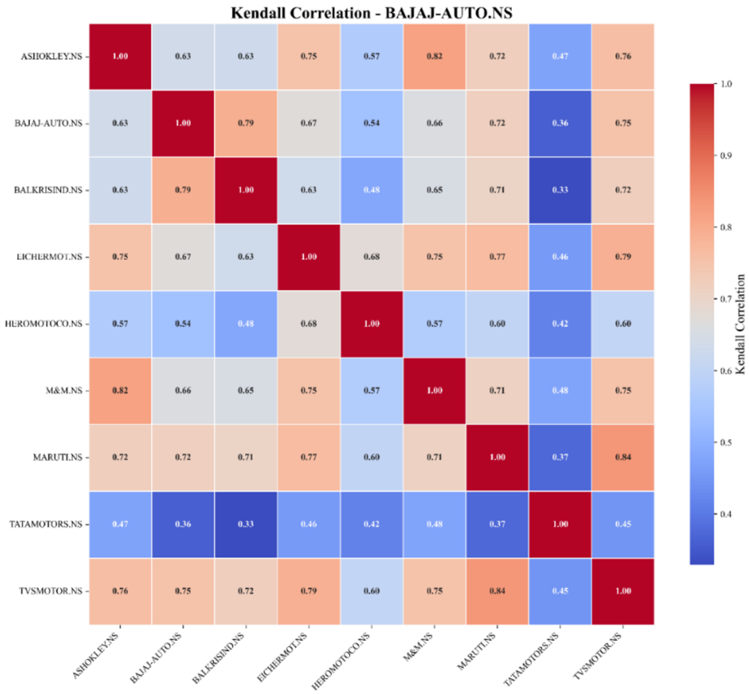

Fig. 9. Kendall Correlation - BAJAJ-AUTO.NS

Interpretation of Figure 9: The Kendall correlation matrix for the Auto sector demonstrates uniformly positive and relatively strong associations, confirming a high degree of monotonic agreement among automotive stocks. This indicates that the rank order of stock returns within the sector tends to move consistently in the same direction.

1. BAJAJ-AUTO.NS exhibits its strongest Kendall relationships with BALKRISIND ($\tau \approx 0.78$), TVSMOTOR ($\tau \approx 0.73$), MARUTI ($\tau \approx 0.68$), and M&M ($\tau \approx 0.65$).

2. Even the comparatively weaker relationships with HEROMOTOCO ($\tau \approx 0.49$) and TATAMOTORS ($\tau \approx 0.44$) remain distinctly positive, reinforcing the overall cohesiveness of the Auto sector. Collectively, these values reflect a tightly coupled rank-ordering structure under Kendall's more stringent measure of concordance.

Pharma Sector Analysis (Centred on GLENMARK.NS)

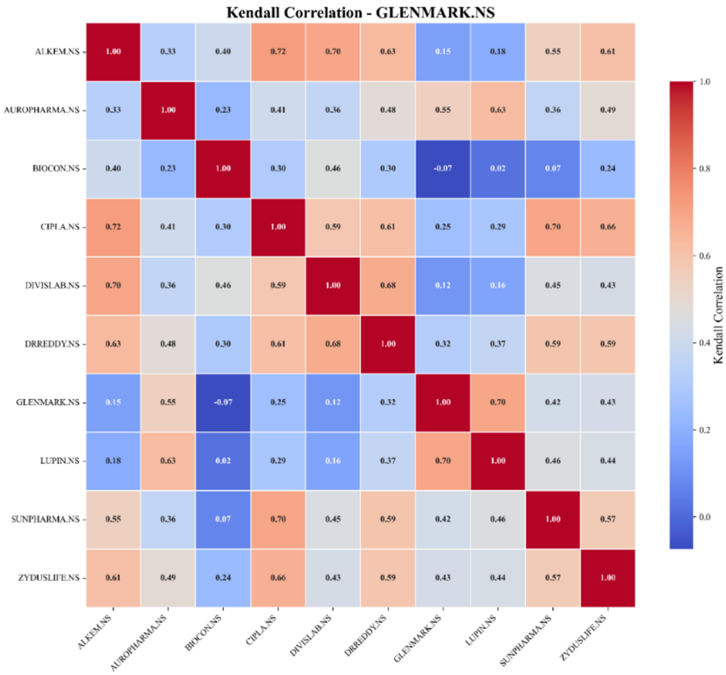

Fig. 10. Kendall Correlation - GLENMARK.NS

Interpretation of Figure 10: In contrast, the Pharma sector (heatmap centred on GLENMARK.NS) displays a far more heterogeneous correlation profile. Unlike the uniformly aligned Auto sector, correlations in Pharma vary widely in strength and even in sign, revealing a fragmented rank-ordering behaviour.

1. GLENMARK.NS exhibits a strong positive relationship with LUPIN ($\tau \approx 0.70$) and moderate associations with AUROPHARMA ($\tau \approx 0.55$), SUNPHARMA ($\tau \approx 0.42$), and ZYDUSLIFE ($\tau \approx 0.43$).

2. However, correlations with CIPLA ($\tau \approx 0.25$), ALKEM ($\tau \approx 0.15$), and DIVISLAB ($\tau \approx 0.12$) are weak, while BIOCON even records a small negative concordance ($\tau \approx -0.07$). This dispersion indicates that rank alignments across the Pharma sector are inconsistent—some firms exhibit synchronous behaviour, while others diverge or move in opposing rank directions.

Cross-Sector Implications

These Kendall results reaffirm the structural divergence between the two sectors: BAJAJ-AUTO.NS operates within a strongly concordant, rank-synchronous environment, whereas GLENMARK.NS is embedded in a fragmented and unevenly correlated landscape.

The contrast between these structures highlights how methodological sensitivity affects perceived co-movement. The GLENMARK–BAJAJ cross-sector pair consistently ranks among those exhibiting the largest differences across Pearson, Spearman, Kendall, and cross-correlation measures.

By linking a company from a tightly coherent rank structure to one within a dispersed and occasionally contradictory network, this pair provides a clear illustration of how sectoral composition and dependency form influence correlation outcomes under different methodological lenses.

3.6 Cross-Correlation Analysis – Temporal Synchronization

While static correlation measures reveal the strength of simultaneous relationships, they cannot capture the time-lagged interactions that characterize complex systems like the stock market. To explore these dynamics, we employ Cross-Correlation Analysis, a method designed to identify how information and momentum propagate over time by measuring the correlation between two return series at various time lags. This temporal analysis uncovers a rich landscape of market behaviors—from immediate synchronization to long-range economic signaling—that are invisible to other methods.

Intra-Sector Dynamics: Distinct Temporal Signatures

Our analysis of the two anchor sectors reveals that even within the Indian market, there is no single temporal pattern. Instead, each sector exhibits a unique "temporal signature" that reflects its underlying structure.

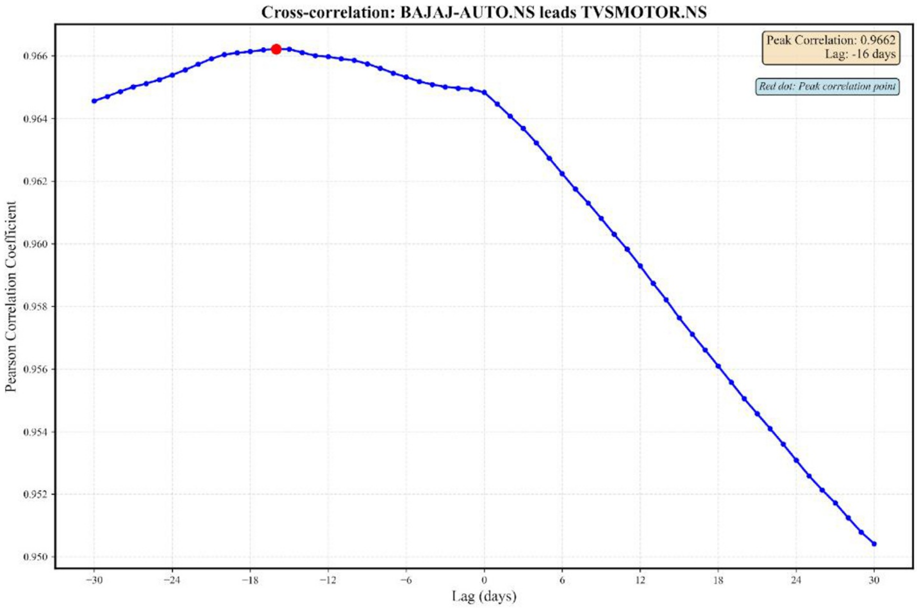

Fig. 11. Cross-correlation: BAJAJ-AUTO.NS leads TVSMOTOR.NS

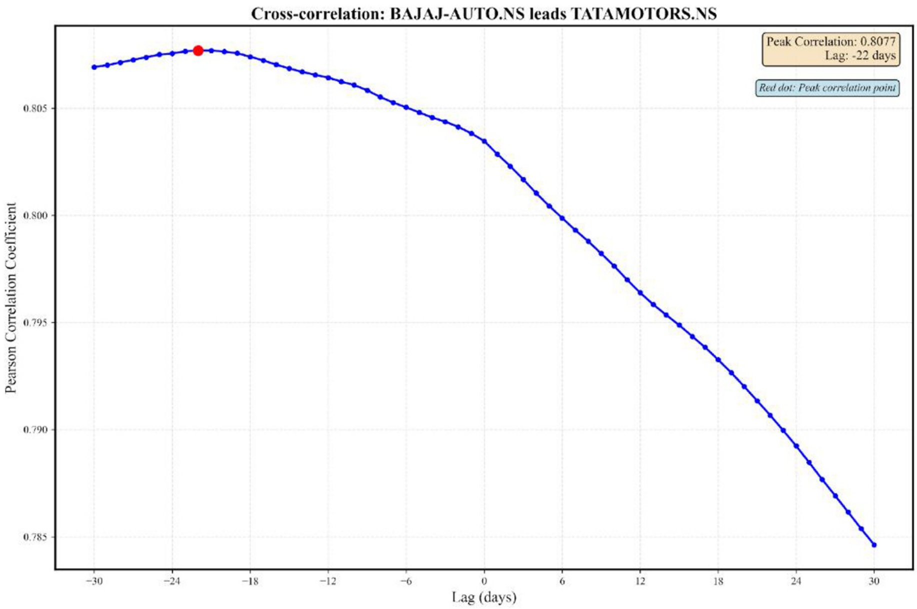

Fig. 12. Cross-correlation: BAJAJ-AUTO.NS leads TATAMOTORS.NS

The Auto Sector: A Pattern of Delayed Synchronicity

Interpretation of Figure 11 and 12: The Auto sector, identified as highly integrated by static measures, surprisingly reveals a consistent pattern of delayed synchronicity. BAJAJ-AUTO.NS emerges as a clear leading indicator, or "bellwether," for its peers, with its market movements preceding the rest of the sector by several weeks.

1. The analysis shows BAJAJ-AUTO.NS leads its strongest peer, TVSMOTOR.NS, by a significant 16 trading days, with a peak correlation of $\rho_t \approx 0.9662$.

2. This leadership effect is even more pronounced with its weakest peer, TATAMOTORS.NS, which it leads by 22 trading days, with a peak correlation of $\rho_t \approx 0.8077$.

This consistent medium-term lag suggests that even in a highly coupled sector, information and sentiment originating from a market leader take considerable time to be fully priced into other constituents, pointing to a potential market inefficiency.

The Pharma Sector: A Mix of Immediate Synchronization and Decoupling

In stark contrast, the fragmented Pharma sector displays a mixed temporal structure characterized by both immediate reactions and complete disconnections.

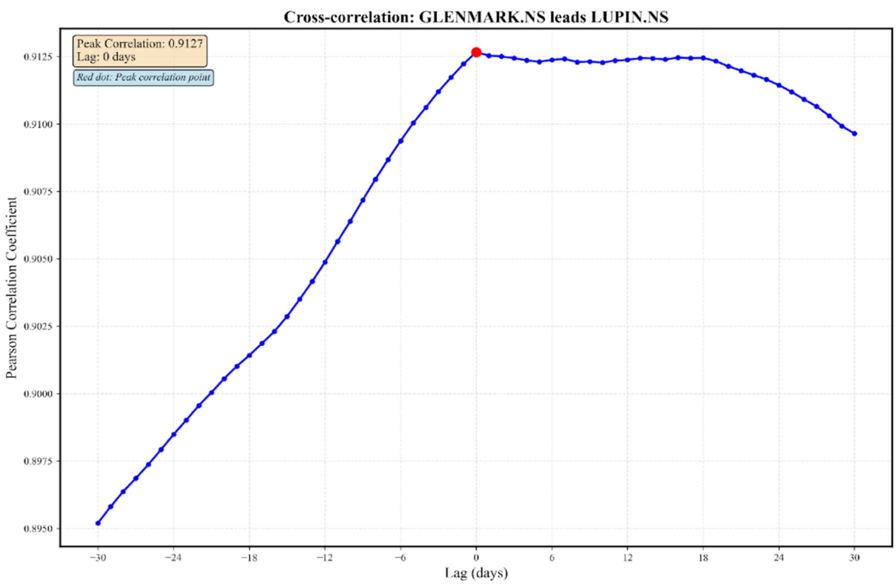

Fig. 13. Cross-correlation: GLENMARK.NS leads LUPIN.NS

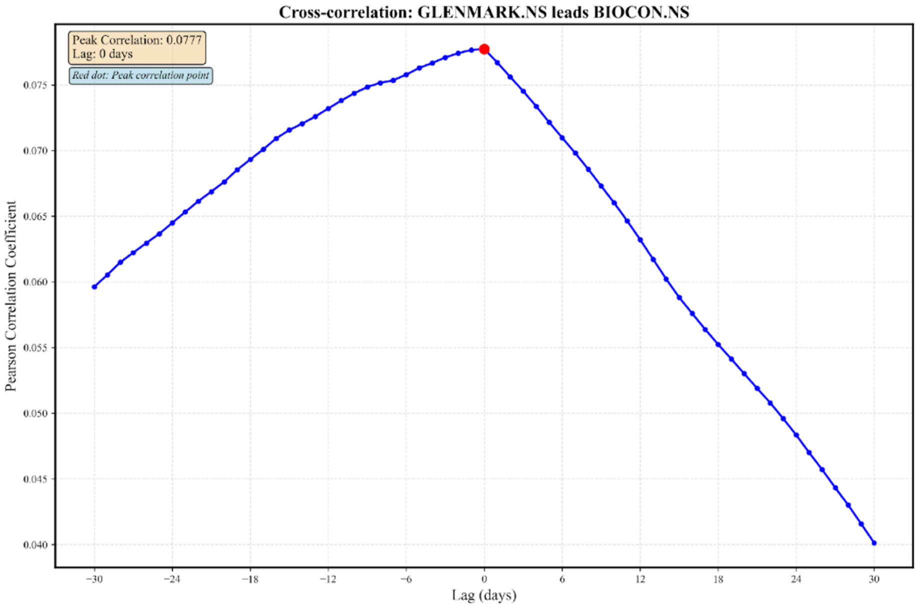

Fig. 14. Cross-correlation: GLENMARK.NS leads BIOCON.NS

Interpretation of Figure 13 and 14:

- **Synchronized Sub-clusters:** The relationship between GLENMARK.NS and its strongest peer, LUPIN.NS, shows a dominant peak correlation of $t \approx 0.9127$ precisely at Lag 0. This signifies immediate, same-day synchronization and suggests that tightly integrated sub-groups within the sector process information efficiently and simultaneously.

- **Decoupled Entities:** Conversely, the analysis between GLENMARK.NS and its weakest peer, BIOCON.NS, shows a negligible correlation ($t \approx 0.0777$) across all time lags, with the insignificant peak also occurring at Lag 0. This temporal decoupling confirms that these companies operate on fundamentally different drivers.

Cross-Sector Dynamics: Cyclical Sectors as Long-Range Economic Indicators

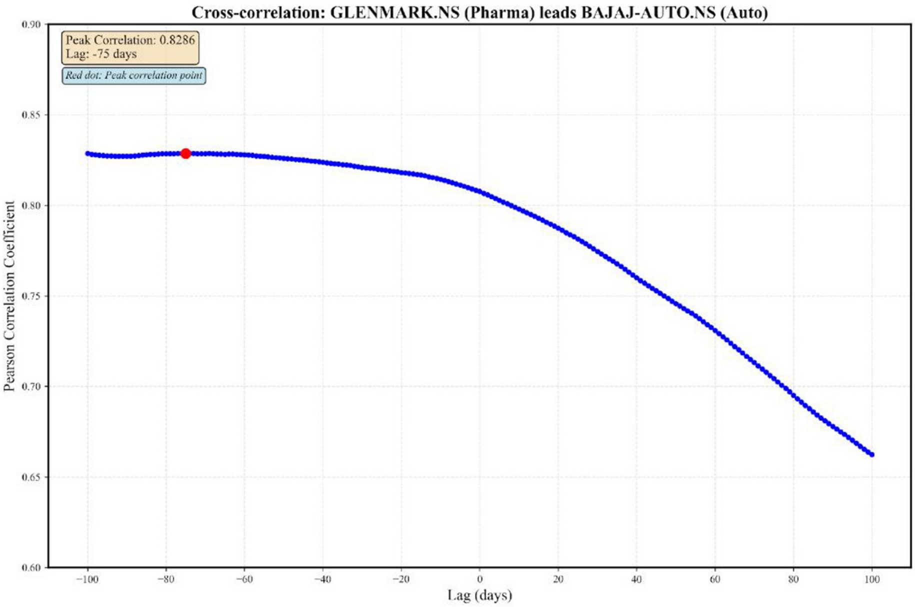

Fig. 15. Cross-correlation: GLENMARK.NS (Pharma) leads BAJAJ-AUTO.NS (Auto)

Interpretation of Figure 15: The most significant finding of this study emerges from the cross-sector analysis between the cyclical BAJAJ-AUTO.NS (Auto) and the defensive GLENMARK.NS (Pharma). This analysis reveals a powerful, long-range predictive signal between two fundamentally different parts of the economy.

Our analysis shows that BAJAJ-AUTO.NS leads GLENMARK.NS by a remarkable 75 trading days, or approximately one full business quarter, with a statistically significant peak correlation of $\rho_t \approx 0.8286$ (Figure 5).

Interpretation

This long-range lead by a cyclical sector over a defensive one is a profound macroeconomic signal. It suggests that shifts in the Auto sector—a key barometer for consumer confidence, economic growth, and "risk-on" investor sentiment—precede corresponding shifts in the more defensive Pharma sector. Positive momentum in the Auto industry appears to be an early indicator of a strengthening economy, the effects of which influence capital flows and sentiment in defensive sectors a quarter later. This finding demonstrates the power of cross-correlation in uncovering complex, long-range economic relationships that are invisible to standard correlation methods and suggests that cyclical sectors can act as leading indicators for defensive ones over a quarterly horizon.

3.7 Comparative Correlation Analysis – Pearson, Spearman, and Kendall Tau

The bar charts visualize the correlation coefficients between the anchor stocks (BAJAJ-AUTO.NS and GLENMARK.NS) and their respective sector peers, comparing the Pearson (linear), Spearman (monotonic), and Kendall Tau (rank-based) methods. This direct comparison highlights the impact of linearity assumptions and ties in the data presented in the previous heatmap analyses.

Auto Sector Comparison (BAJAJ-AUTO.NS)

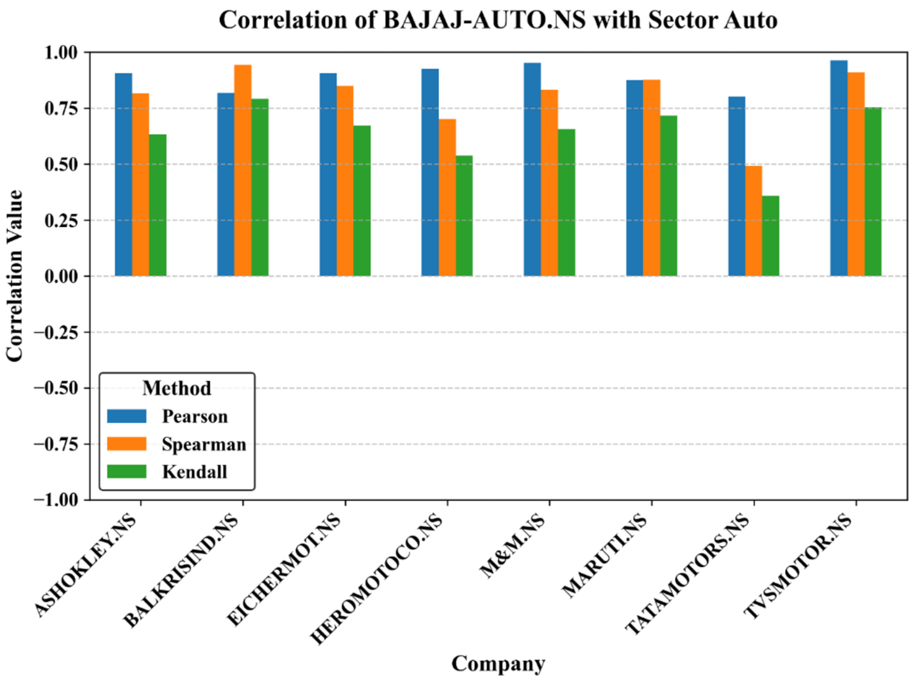

Fig. 16. Correlation of BAJAJ-AUTO.NS with Sector Auto

Interpretation of Figure 16: The correlation of BAJAJ-AUTO.NS with its Auto sector peers is uniformly high and positive across all three methods, reinforcing the sector's integrated structure. BAJAJ-AUTO.NS exhibits its highest co-movement with TVSMOTOR ($\rho \approx 0.96$) and M&M ($\rho \approx 0.95$), followed closely by HEROMOTOCO ($\rho \approx 0.92$), EICHERMOT ($\rho \approx 0.90$), and MARUTI ($\rho \approx 0.88$).

1. Uniformly High Correlation: For nearly all pairs, the coefficients for all three methods cluster closely, typically ranging between 0.6 and 1.0. This close clustering demonstrates that the relationships in the Auto sector are not only strong and monotonic (Spearman, Kendall) but also highly linear (Pearson).

2. Method Hierarchy: A clear hierarchy is consistently observed: Pearson Spearman Kendall.

   a) Pearson coefficients are generally the highest, indicating a strong degree of linear proportionality in the returns.
   b) Kendall Tau coefficients are consistently the lowest, which is expected as Kendall Tau is a more conservative and robust measure of rank concordance compared to Spearman.

3. Structural Consistency: The minimal spread between the three bars for most companies (e.g., EICHERMOT.NS, M&M.NS, MARUTI.NS) implies that the strong dependence structure in the Auto sector is robust, regardless of whether the co-movement is measured by linear fit or by rank ordering. The only pair where the measures show a slightly larger gap is with TATAMOTORS.NS, where the correlation is the lowest across the sector, suggesting a somewhat weaker or less perfectly linear dependence.

Pharma Sector Analysis (Centred on GLENMARK.NS)

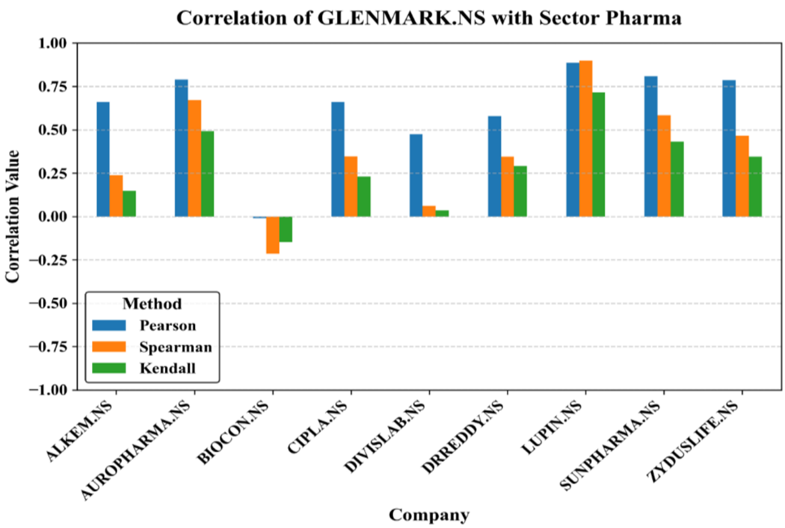

Fig. 17. Correlation of GLENMARK.NS with Sector Pharma

Interpretation of Figure 17: The correlation of GLENMARK.NS with its Pharma sector peers is highly heterogeneous, with significant variation in coefficients and method discrepancies across different pairs.

1. **Positive Correlations:** For most peers (e.g., AUROPHARMA.NS, LUPIN.NS, SUNPHARMA.NS, ZYDUSLIFE.NS), the correlations are positive, with LUPIN.NS exhibiting the strongest relationship across all methods (>0.7).

2. **The BIOCON.NS Anomaly:** The pair with BIOCON.NS clearly stands out.
   a) Pearson is near zero ((≈0.08), indicating no linear relationship.
   b) Spearman is negative (≈0.12), and Kendall is negative (≈-0.09), suggesting a slight inverse relationship in the ranking of returns. This contrast highlights that even when linearity is absent, a slight non-linear or rank-based inverse co-movement can exist, a key indicator of sector fragmentation.

3. **Divergence in Weaker Pairs:** For pairs with weaker overall relationships (e.g., ALKEM.NS, DIVISLAB.NS), the correlation values across the three methods show a greater divergence. For ALKEM.NS, the Pearson is high (≈0.65), while the Spearman (≈0.24) and Kendall (≈0.15) are significantly lower. This indicates that while a relatively strong linear relationship may exist, the rank concordance is much weaker, suggesting the relationship is dominated by non-monotonic, non-linear factors that weaken rank-based co-movement.

3.8 Time Series Validation

To visually validate the Pearson correlation coefficients identified in Section 3.3, we compare the Closing Price (INR) movements of the anchor stocks (BAJAJ-AUTO.NS and GLENMARK.NS) against their strongest and least synchronous correlated peers within their respective sectors. The analysis of these time-series charts provides direct visual evidence for the distinct linear dependence structures of the Auto and Pharma sectors.

Auto Sector: BAJAJ-AUTO.NS Performance Analysis

Fig. 18. Price Movement Comparison in Auto Sector (Highlighting BAJAJ-AUTO.NS Performance)

Interpretation of Figure 18: The price movement chart for the Auto sector (BAJAJ-AUTO.NS, TVSMOTOR.NS, and TATAMOTORS.NS) highlights the difference between highly integrated and least synchronous pairs centered on BAJAJ-AUTO.NS.

1. **Highest Correlated Peer (TVSMOTOR.NS):** The price movement of BAJAJ-AUTO.NS and TVSMOTOR.NS exhibits a remarkable degree of synchronization in price action. The two lines consistently follow the same trend, particularly during major inflection points, such as the steep rally that began in 2023. While the absolute price levels are different (BAJAJ-AUTO.NS at ₹ 8702 vs. TVSMOTOR.NS at ₹ 3414 at the end of the period), the timing and proportional severity of their movements are highly parallel. This visual evidence directly validates the strong Pearson correlation (($\rho \approx 0.96$), confirming a tight linear co-movement in the magnitude of returns.

2. **Weakest Correlated Peer (TATAMOTORS.NS):** The closing price path of TATAMOTORS.NS (₹ 673) shows a comparative decoupling from the anchor stock BAJAJ-AUTO.NS. While its Pearson correlation remains a high $\rho \approx 0.80$, its lowest in the sector, the chart shows that the stock's major price movements are less synchronous and proportionally distinct from BAJAJ-AUTO.NS. This confirms that even within a highly coupled segment, linear divergence can exist, indicating TATAMOTORS.NS is driven by relatively more idiosyncratic factors compared to the rest of the auto index.

Pharma Sector: GLENMARK.NS Performance Analysis

Fig. 19. Price Movement Comparison in Pharma Sector (Highlighting GLENMARK.NS Performance)

Interpretation of Figure 19: The price movement chart for the Pharma sector (GLENMARK.NS, LUPIN.NS, and BIOCON.NS) demonstrates the sector's high heterogeneity, clearly showing the contrast between a strong positive linear relationship and an almost zero-correlation relationship.

1. **Highest Correlated Peer (LUPIN.NS):** The price path of GLENMARK.NS (₹ 1972) and LUPIN.NS (₹ 1920) shows a substantial degree of alignment in price levels and trajectory, especially from 2023 onwards. The two lines track each other closely in terms of direction and timing of major rallies and corrections. This parallel movement validates the strong Pearson correlation ($\rho \approx 0.91$) between the two stocks, confirming that they share a common linear risk exposure or operational driver within the Pharma segment.

2. **Weakest Correlated Peer (BIOCON.NS):** The price path of BIOCON.NS (₹ 339) stands in dramatic contrast to the movement of GLENMARK.NS and LUPIN.NS. BIOCON.NS shows a history of large, volatile, and non-synchronous movements at a significantly lower price base. There are long periods where the price movements are entirely uncoupled from the anchor stocks. This visual chaos and lack of coordinated movement provide compelling evidence for the near-zero Pearson correlation ($\rho \approx 0.08$), confirming that the stock returns are driven by highly idiosyncratic, company-specific factors that negate any meaningful linear co-movement with GLENMARK.NS.

4 Conclusion

This study demonstrates that inter-sectoral dependence in the Indian stock market is fundamentally multi-dimensional and cannot be reliably represented using a single correlation measure. By jointly analyzing linear (Pearson), rank-based (Spearman and Kendall), and temporal (cross-correlation) dependencies, the proposed framework advances correlation analysis toward a structured, quantitative assessment of dependence heterogeneity.

The primary methodological contribution is the introduction of the Average Absolute Deviation (AvgDiff) metric, which numerically captures divergence across dependence measures. Cross-sector stock pairs exhibit AvgDiff values in the range 0.35–0.50, providing direct evidence that linear, monotonic, and temporal dependencies differ substantially across sectoral boundaries. This result confirms the limitations of single-metric correlation analysis in emerging markets.

Empirically, the Auto sector displays consistently high dependence across all measures, with Pearson and Spearman coefficients exceeding 0.85 for most intra-sector pairs. In contrast, the Pharma sector shows wide dispersion, with Pearson correlations as low as 0.08 and rank-based coefficients extending into negative values, indicating heterogeneous inter-stock relationships. These quantified differences reveal the coexistence of highly synchronized and weakly coupled dependence structures within the same market.

Temporal analysis further strengthens the findings by identifying long-horizon lead–lag effects, including a 75-trading-day lag between representative Auto and Pharma stocks. Such delayed interactions are not observable through static correlations and highlight the importance of temporal dependence in cross-sector analysis.
Overall, the multi-metric framework transforms correlation analysis into a diagnostic methodology for identifying structural and temporal dependence patterns. The results offer actionable insights for systemic risk assessment, diversification strategies, and time-aware portfolio design, while providing a scalable foundation for future research on dynamic financial networks and predictive dependence modeling.

References

1. Sharma, C., Banerjee, K.: A study of correlations in the stock market. Physica A: Statistical Mechanics and Its Applications 432, 321–330 (2015)

2. Kulkarni, V., Deo, N.: Correlation and volatility in an Indian stock market: A random matrix approach. The European Physical Journal B 60(1), 101–109 (2007)

3. Essam, F., El Hashash, Ali, S.R.H.: A comparison of the Pearson, Spearman rank and Kendall tau correlation coefficients using quantitative variables. Asian Journal of Probability and Statistics 20(3), 36–48 (2022)

4. Puth, M.-T., Neuhäuser, M., Ruxton, G.D.: Effective use of Spearman's and Kendall's correlation coefficients for association between two measured traits. Animal Behaviour 102, 77–84 (2015)

5. Da Costa Lewis, N.: Correlation analysis. In: Energy risk modeling: Applied modeling methods for risk managers, pp. 125–150. Palgrave Macmillan, (2005)

6. Eftekharypour, Y., Ngo, C.H., Ong, H.: Partial correlation threshold network analysis of Malaysia stock market. In: 2018 4th International Conference on Computer and Information Sciences (ICCOINS), pp. 1–6. IEEE, (2018)

7. Bakar, N.A., Rosbi, S.: Statistical non-parametric correlation diagnostic for dynamic behavior of Malaysian currency with world crude oil price. International Journal of Economics, Commerce and Management 5(5), 1–13 (2017)

8. Kshirsagar, S., Thombre, S.: Exploring ML techniques for correlating sentiment with stock prices in India. In: 2025 1st International Conference on AIML-Applications for Engineering & Technology (ICAET), pp. IEEE, (2025)

9. Az-Zahra, F., et al.: The correlation analysis of stock price towards public sentiment in automotive industry. In: 2021 International Conference on Data Science and Its Applications (ICoDSA), pp. 240–244. IEEE, (2021)

10. Das, T., Halder, A., Saha, G.: Application of density-based clustering approaches for stock market analysis. Applied Artificial Intelligence 38(1), (2024)

11. Sun, J., Sheng, H.: Multifractal detrended cross-correlation analysis of Chinese stocks. In: 2010 Third International Conference on Business Intelligence and Financial Engineering, pp. 118–122. IEEE, (2010)

12. Yi, Z.: Optimized portfolio strategy in Chinese growth enterprise market: Based on the detrended cross-correlation analysis method. In: 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME), pp. 103–107. IEEE, (2020)