




NETWORK ANALYSIS

Indian Stock Market

ECS-414



Problem Statement

Creation and analysis of a stock market network using the data from Indian Stock market.

The objective here is to study correlation between stocks traded at the National Stock Exchange (NSE) between the period of 1st March, 2020 to 1st March, 2024. The stocks chosen were on the basis of the market capitalization data published by NSE¹.

The Equity (EQ) series has been chosen with the Normal segment (N) as the market type.

¹ [All Companies based on Market Capitalisation \(December 31, 2023\)](#)

Literature Review

Network analysis of a financial market based on genuine correlation and threshold method

By A. Namaki, A.H. Shirazi, R. Raei, G.R. Jafari

A network perspective of the stock market

By Chi K. Tse, Jing Liu, Francis C.M. Lau

Complex networks analysis in Iran stock market: The application of centrality

By Hadi Esmaeilpour Moghadam, Teymour Mohammadi, Mohammad Fegghi Kashani, Abbas Shakeri

Configuration

Number of nodes	200
Threshold	0.45
Number of edges	650
Diameter of Giant Component	10
Edge Density	0.03
Average Clustering Coefficient	0.399437
Degree Assortativity	0.080832

Methodology

- Data collection
- Data processing
- Calculation of correlation
- Threshold Selection
- Network Formation
- Network Analysis

Data Collection

The data is collected from NSE's website by means of the *nsepython* library which communicates with their REST APIs.

Through this package we communicate with the NSE's APIs and save the data for the stocks of interest as separate CSVs.



Data Processing

- Get rid of the useless metadata.
- Check for ordering among the data.
- Remove redundant data points and duplicates.
- Save data to a new folder.

Correlation Calculation

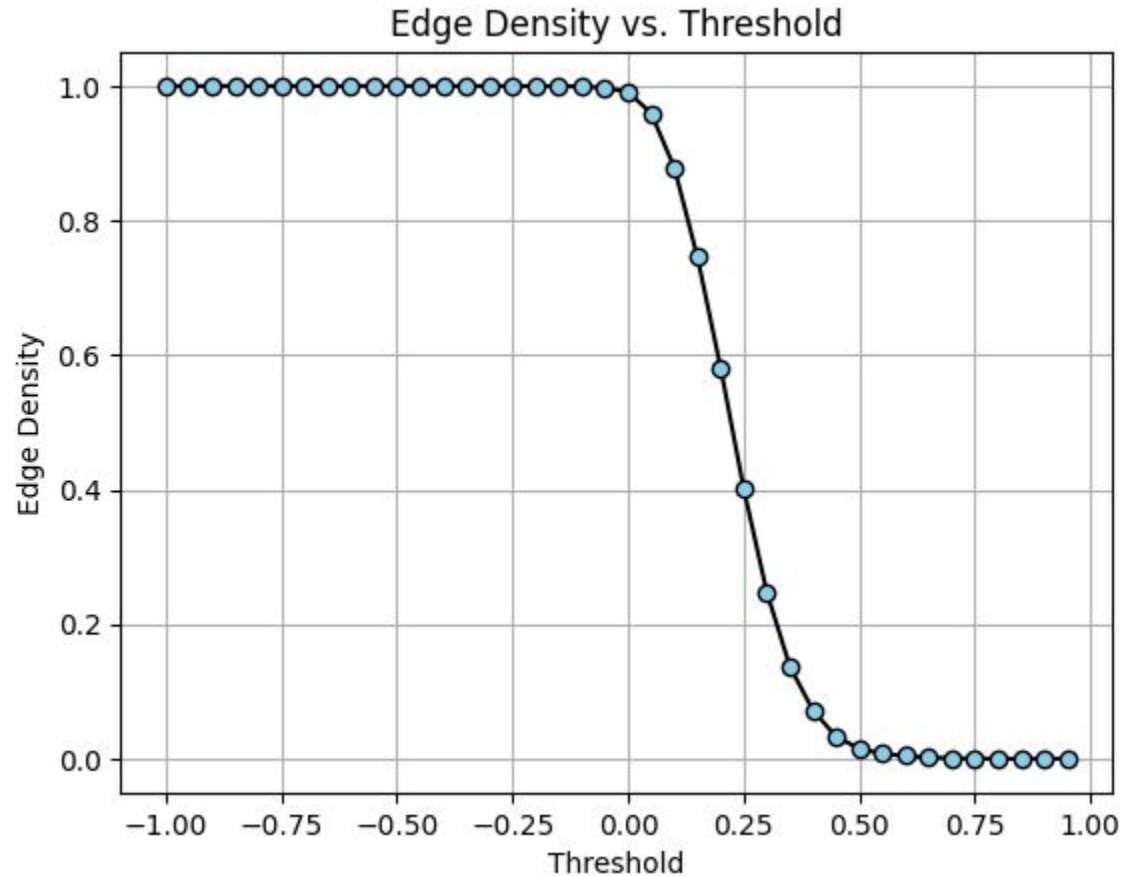
- Calculated the log difference of the “closing price” of stocks for each day

$$r_i(t) = \ln p_i(t) - \ln p_i(t - 1)$$

- Calculated the correlation of all of the stocks with each other

$$c_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2) (\langle r_j^2 \rangle - \langle r_j \rangle^2)}}$$

Threshold Selection



it can be seen that the number of edges in the stock correlation network decreases as the threshold value θ increases.

Threshold Method

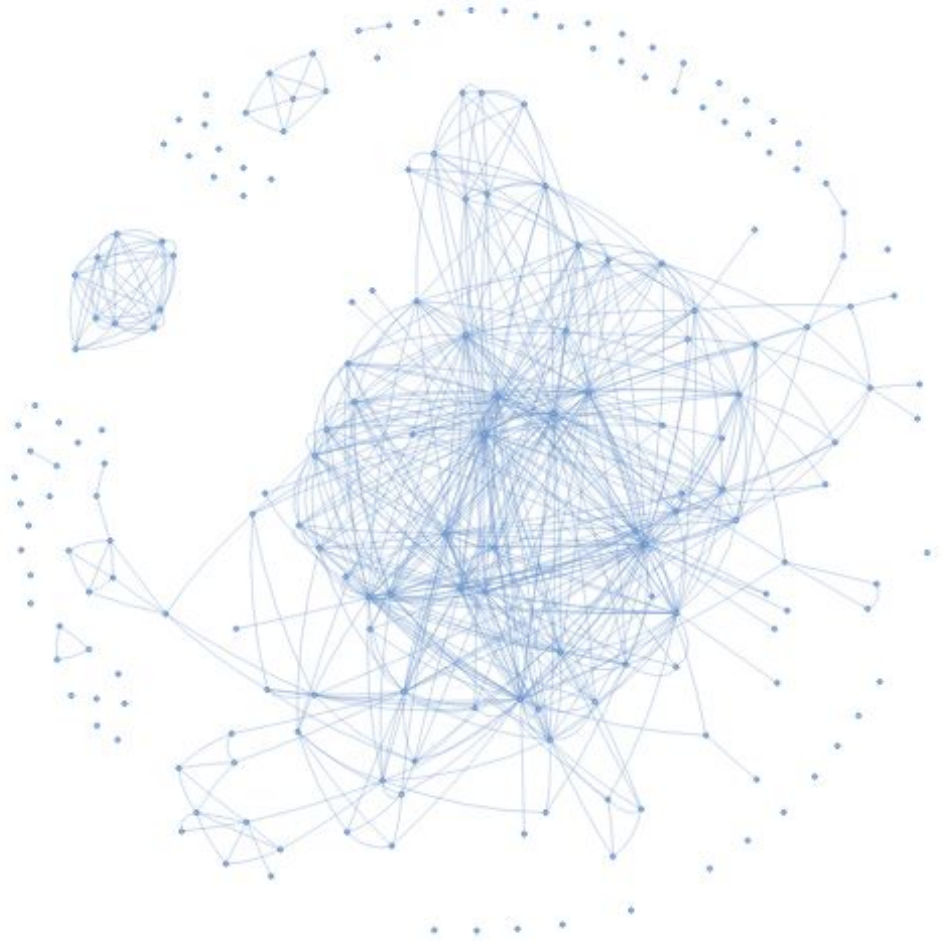
If the correlation coefficient c_{ij} is greater than or equal to θ , we add an undirected edge connecting the vertices i and j .

Θ : Threshold

The selected threshold was 0.45.

Network Formation

Number of nodes = 200
Number of edges = 650
Edge density = 0.03
Diameter of giant component = 10
Mean degree of network = 6.5



Network Analysis

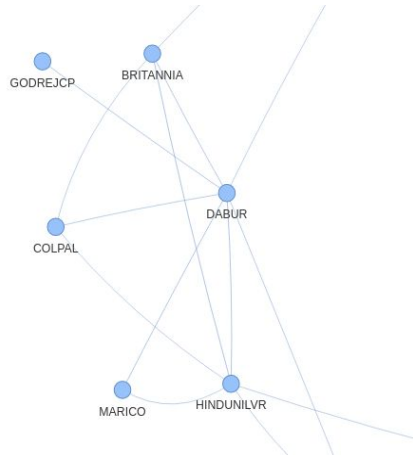
- Communities
- Degree Distribution
- Degree Centrality
- Pagerank Centrality
- Eigenvector Centrality
- Betweenness Centrality
- Closeness Centrality

Communities formed

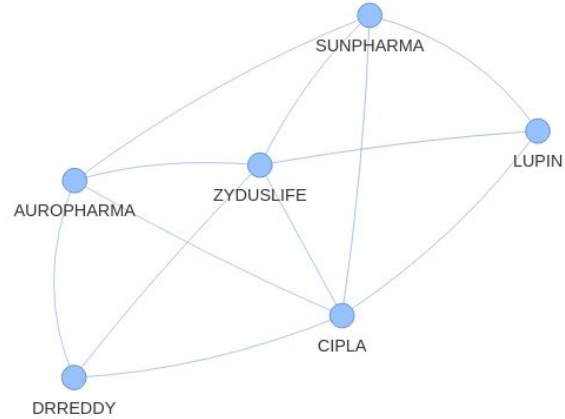
We use the Clauset-Newman-Moore greedy modularity maximization to find the community partition with the largest modularity.

- CANBK KOTAKBANK RECLTD INDUSINDBK FEDERALBNK BANDHANBNK UPL DLF TRENT HEROMOTOCO SHIRAMFIN TATAMOTORS IDFCFIRSTB BHARATFORG GODREJPROP SONACOMS BAJAJ-AUTO RELIANCE TITAN ABCAPITAL M&MFIN MOTHERSON L&TFH OBEROIRLTY ASHOKLEY BAJFINANCE CHOLAFIN SBICARD TVSMOTOR BOSCHLTD M&M AXISBANK SIEMENS CONCOR HAVELLS SBIN GMRINFRA DEEPAKNTR PRESTIGE HDFCAMC BHARTIARTL JSWSTEEL GRASIM MUTHOOTFIN LT PFC INDHOTEL HDFCBANK MARUTI ICICIBANK
- OIL IOC JINDALSTEL TATAPOWER SAIL HINDPETRO NHPC POWERGRID HINDALCO VEDL SJVN BHEL COALINDIA NMDC BPCL NLCINDIA JSL TORNTPOWER NTPC GAIL ONGC
- MRF BERGEPAINT TATACONSUM BRITANNIA SBILIFE HINDUNILVR COLPAL GODREJCP BALKRISIND MARICO HDFCLIFE ASIANPAINT PIDILITIND ICICIGI ICICIPRULI DABUR ULTRACEMCO
- DALBHARAT JIOFIN SHREECEM ATGL ADANIENIT ADANIGREEN ACC AWL ADANIENSOL ADANIPOWER AMBUJACEM
- HCLTECH PERSISTENT LTTS LTIM INFY WIPRO TECHM COFORGE TCS MPHASIS
- CENTRALBK UNIONBANK INDIANB IDBI BANKBARODA PNB BANKINDIA IOB UCOBANK
- ZYDUSLIFE SUNPHARMA LUPIN CIPLA AUOPHARMA DRREDDY
- NIACL GICRE LIC
- ASTRAL DIXON
- INDUSTOWER IDEA

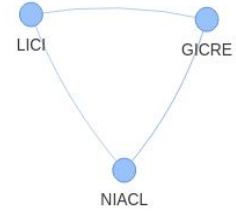
Communities



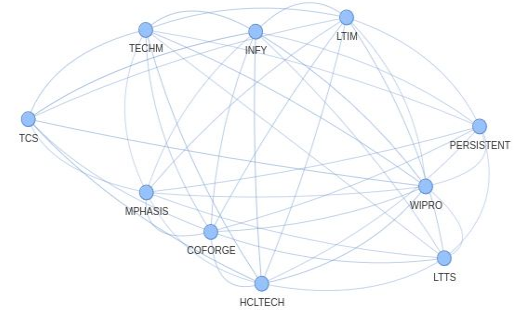
Consumer



Pharmaceutical

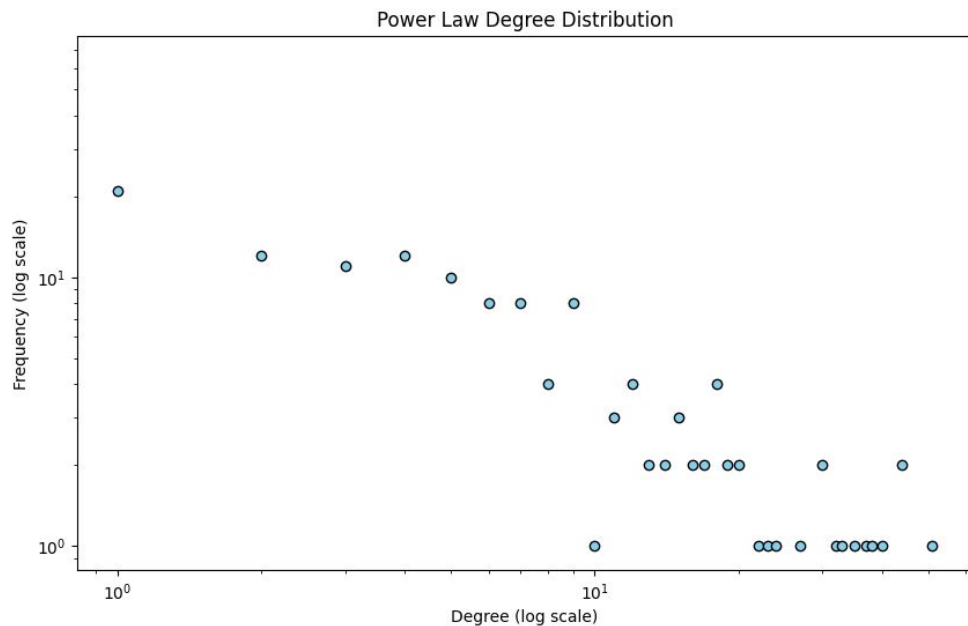
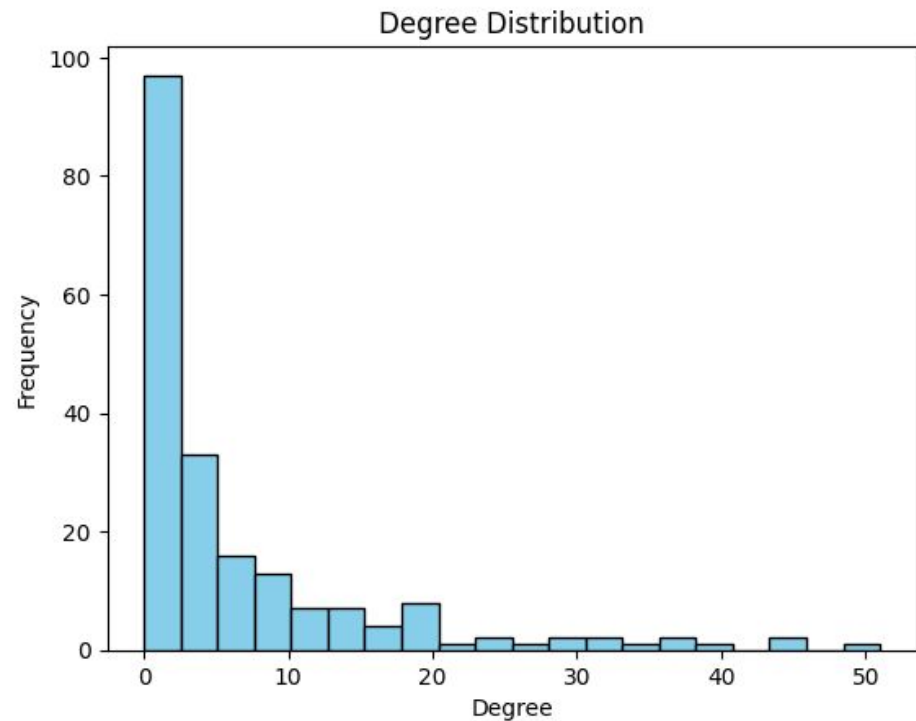


Insurance



Tech

Degree Distribution



Degree Centrality

Top 10 central nodes based on degree centrality

$$C_D(i) = \frac{d(i)}{n - 1}$$

L&TFH	0.2563
FEDERALBNK	0.2211
DLF	0.2211
SBIN	0.2010
IDFCFIRSTB	0.1910
GRASIM	0.1859
CANBK	0.1759
BAJFINANCE	0.1658
LT	0.1608
ICICIBANK	0.1508

Page-Rank Centrality

Top 10 central nodes based on pagerank centrality

$$\mathbf{x} = (\mathbf{I} - \alpha \mathbf{A} \mathbf{D}^{-1})^{-1} \mathbf{1}.$$

L&TFH	0.0277
DLF	0.0246
FEDERALBNK	0.0223
GRASIM	0.0207
SBIN	0.0205
IDFCFIRSTB	0.0188
CANBK	0.0177
BAJFINANCE	0.0172
ICICIBANK	0.0154
LT	0.0151

Eigenvector Centrality

Top 10 central nodes based on eigenvector centrality

$$x_i = \kappa^{-1} \sum_{j=1}^n A_{ij} x_j.$$

L&TFH	0.2611
FEDERALBNK	0.2529
SBIN	0.2406
DLF	0.2334
IDFCFIRSTB	0.2227
CANBK	0.2119
LT	0.2092
AXISBANK	0.2056
BAJFINANCE	0.2054
ICICIBANK	0.2025

Betweenness Centrality

Top 10 central nodes based on betweenness centrality

$$x_i = \sum_{st} \frac{n_{st}^i}{g_{st}},$$

L&TFH	0.052
DLF	0.0391
GRASIM	0.0391
ADANIENT	0.0317
ULTRACEMCO	0.0305
TATACONSUM	0.0292
SBIN	0.0218
IDFCFIRSTB	0.0185
COALINDIA	0.0181
FEDERALBNK	0.0176

Closeness centrality

Top 10 central nodes based on closeness centrality

$$CC(i) = \frac{N - 1}{\sum_j d(i, j)}$$

$d(i, j)$ is the length of the shortest path between nodes i and j

L&TFH	0.3102
DLF	0.2909
SBIN	0.2895
GRASIM	0.2855
FEDERALBNK	0.2828
CANBK	0.2714
IDFCFIRSTB	0.2714
BAJFINANCE	0.2644
ULTRACEMCO	0.2644
LT	0.2610

References

- Esmailpour Moghadam, Hadi, Teymour Mohammadi, Mohammad Fegghi Kashani, and Abbas Shakeri. “Complex Networks Analysis in Iran Stock Market: The Application of Centrality.” *Physica A: Statistical Mechanics and Its Applications* 531 (October 1, 2019): 121800. <https://doi.org/10.1016/j.physa.2019.121800>.
- Hagberg, Aric, Pieter J. Swart, and Daniel A. Schult. “Exploring Network Structure, Dynamics, and Function Using NetworkX.” Los Alamos National Laboratory (LANL), Los Alamos, NM (United States), January 1, 2008. <https://www.osti.gov/biblio/960616>.
- Tse, Chi K., Jing Liu, and Francis C. M. Lau. “A Network Perspective of the Stock Market.” *Journal of Empirical Finance* 17, no. 4 (September 1, 2010): 659–67. <https://doi.org/10.1016/j.jempfin.2010.04.008>.
- Namaki, A., A. H. Shirazi, R. Raei, and G. R. Jafari. “Network Analysis of a Financial Market Based on Genuine Correlation and Threshold Method.” *Physica A: Statistical Mechanics and Its Applications* 390, no. 21 (October 15, 2011): 3835–41. <https://doi.org/10.1016/j.physa.2011.06.033>.
- Clauset, Aaron, M. E. J. Newman, and Christopher Moore. “Finding Community Structure in Very Large Networks.” *Physical Review E* 70, no. 6 (December 6, 2004): 066111. <https://doi.org/10.1103/PhysRevE.70.066111>.
- Newman, Mark. *Networks: An Introduction*. USA: Oxford University Press, Inc., 2018.