Network Analysis of Indian Stock Market

Ananya Gandhi (20319)
Janhavi Hatwar (19138)
Mohammad Anish (19195)

Introduction

1. In this project, we perform a network analysis of stock price data from the Indian stock market. The objective is to explore the interconnections between different stocks based on their price movements over time.
2. By constructing a network of stocks and analyzing its properties, we aim to gain insights into the underlying structure and dynamics of the stock market.
3. The code performs a comprehensive network analysis on the dataset, utilizing various libraries such as pandas, numpy, matplotlib, seaborn, and networkx to process the data, calculate correlations, and visualize the network.
4. It analyzes the relationships and structure of the stock market by constructing a network based on the correlations between stock return prices.

Data Collection and Preprocessing

- We collected historical stock price data for a variety of stocks listed on the Indian stock market.
- The data was obtained from the **National Stock Exchange (NSE)** and consisted of **daily closing prices** for each stock over a specified time period **(01 Feb, 2021 to 31 Jan, 2022).**
- Our code reads the list of files in the dataset directory, sorts them, and removes the first file (which is an index).

Correlation Calculation

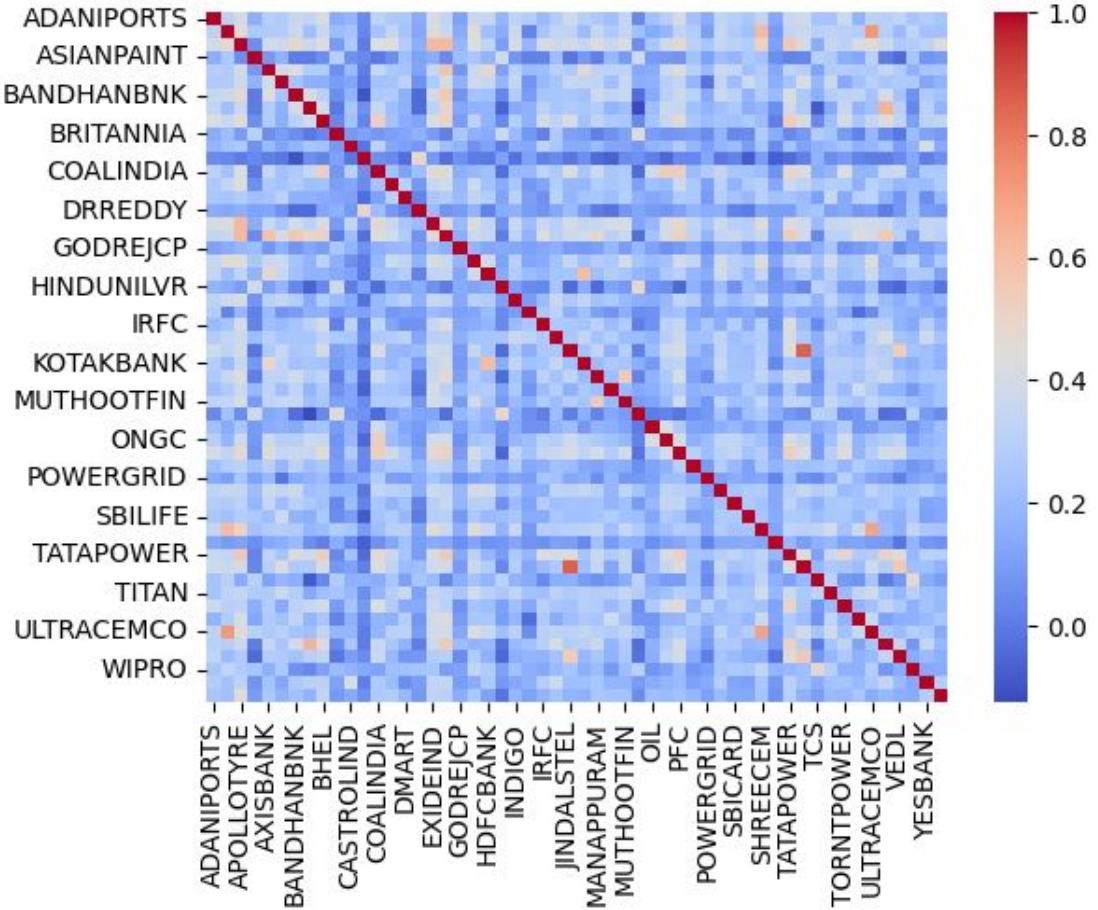
- Our code then calculates the correlation matrix between the return prices of different stocks.
- The correlation matrix is then calculated using the **Pearson correlation coefficient method** and printed.
- After calculating the Pearson correlation coefficient between the log returns of each pair of stocks; ***This coefficient measures the strength of the linear relationship between the variables.***

Dataset

	Date	series	OPEN	HIGH	LOW	PREV. CLC	ltp	close	vwap	52W H	52W L	VOLUME	VALUE	No of trade	CLOSE
0	01-Feb-21	EQ	514.3	545.45	500.1	509.7	543.8	542.75	521.81	562.5	203	10259084	5,35,32,54,917.15	144291	542.75
1	02-Feb-21	EQ	548.7	554.85	540	542.75	553.3	550.8	547.68	562.5	203	7160471	3,92,16,59,690.50	72629	550.8
2	03-Feb-21	EQ	549.55	562.3	542	550.8	560	560.1	555.66	562.5	203	7462289	4,14,65,13,325.30	73872	560.1
3	04-Feb-21	EQ	557.9	573.4	554.05	560.1	568.7	568.4	565.59	573.4	203	6850462	3,87,45,50,036.00	76877	568.4
4	05-Feb-21	EQ	571	573.95	561.65	568.4	567.6	566.95	568.7	573.95	203	5586868	3,17,72,37,232.50	65457	566.95
5	08-Feb-21	EQ	572.15	594	571.1	566.95	583.3	582.05	583.52	594	203	8887710	5,18,61,33,387.80	118318	582.05
6	09-Feb-21	EQ	589.65	589.8	573.2	582.05	582.2	581.95	581.89	594	203	10917915	6,35,30,78,363.25	110355	581.95
7	10-Feb-21	EQ	584.75	587	550.5	581.95	578.4	576.85	567.22	594	203	19987392	11,33,72,93,385.70	202060	576.85
8	11-Feb-21	EQ	576.85	593.7	576.3	576.85	590.25	591.85	585.92	594	203	10595917	6,20,83,10,908.15	98093	591.85
9	12-Feb-21	EQ	592.3	621.65	586.1	591.85	611.75	616.6	606.51	621.65	203	24953701	15,13,46,77,388.95	253089	616.6
10	15-Feb-21	EQ	617.95	639.9	610.8	616.6	625.8	627.75	627.65	639.9	203	15146033	9,50,64,62,378.65	184606	627.75
11	16-Feb-21	EQ	629.8	643.5	628.05	627.75	637.4	636.55	637.44	643.5	203	15120148	9,63,81,25,111.15	241095	636.55
12	17-Feb-21	EQ	658	670	646.9	636.55	654	655.75	658.2	670	203	34868085	22,95,01,89,490.75	469652	655.75
13	18-Feb-21	EQ	660.8	670	649.05	655.75	664.2	664.05	659.78	670	203	13189027	8,70,18,91,515.10	161477	664.05
14	19-Feb-21	EQ	664.95	676.65	628.65	664.05	655	653.5	658.84	676.65	203	17553743	11,56,50,22,396.50	219949	653.5
15	22-Feb-21	EQ	658.8	678.5	649	653.5	672	673.8	665.66	678.5	203	16739107	11,14,24,95,641.05	200480	673.8
16	23-Feb-21	EQ	680	682.7	659.25	673.8	662	662	668.7	682.7	203	11247164	7,52,09,40,821.50	160718	662
17	24-Feb-21	EQ	666.7	671	653.4	662	666.4	665.2	665.04	682.7	203	3919187	2,60,63,99,185.80	57372	665.2
18	25-Feb-21	EQ	669.4	711.35	667.15	665.2	705	704.3	695.14	711.35	203	18596690	12,92,72,21,549.30	226639	704.3
19	26-Feb-21	EQ	697.1	703.45	668	704.3	668.25	675.9	686.34	711.35	203	19909990	13,66,51,13,847.35	220978	675.9
20	01-Mar-21	EQ	684.6	704.65	673.3	675.9	697.6	692.9	688.54	711.35	203	11600468	7,98,74,21,582.95	179958	692.9
21	02-Mar-21	EQ	697.6	724.7	696	692.9	721.25	720.5	708.76	724.7	203	11486477	8,14,11,58,368.05	170602	720.5
22	03-Mar-21	EQ	732	754.95	723.4	720.5	733.25	729.85	739.56	754.95	203	19961472	14,76,26,53,883.70	285506	729.85
23	04-Mar-21	EQ	733.25	765.7	715.85	729.85	751.55	752.45	749.5	765.7	203	30607959	22,94,05,36,363.40	388068	752.45
24	05-Mar-21	EQ	752.45	761	730.5	752.45	749.65	748.95	748.31	765.7	203	17665200	13,21,89,90,805.30	245042	748.95

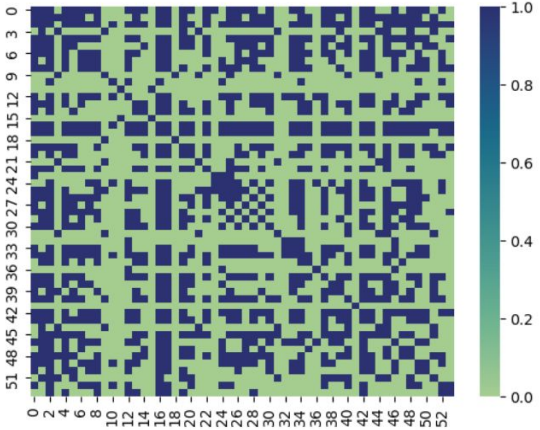
Index	Company	Index	Company	Index	Company	Index	Company
0	ADANIPORTS	14	DMART	28	MANAPPURAM	42	TATAPOWER
1	AMBUJACEM	15	DRREDDY	29	MARUTI	43	TATASTEEL
2	APOLLOTYRE	16	EXIDEIND	30	MUTHOOTFIN	44	TCS
3	ASIANPAINT	17	FEDERALBNK	31	NESTLEIND	45	TITAN
4	AXISBANK	18	GODREJCP	32	OIL	46	TORNTPOWER
5	BAJAJFINSV	19	HAVELLS	33	ONGC	47	TVSMOTOR
6	BANDHANBNK	20	HDFCBANK	34	PFC	48	ULTRACEMCO
7	BANKINDIA	21	HINDUNILVR	35	POLYCAB	49	UNIONBANK
8	BHEL	22	INDIGO	36	POWERGRID	50	VEDL
9	BRITANNIA	23	IRCTC	37	RELIANCE	51	WIPRO
10	CASTROLIND	24	IRFC	38	SBICARD	52	YESBANK
11	CIPLA	25	ITC	39	SBILIFE	53	ZEEL
12	COALINDIA	26	JINDALSTEL	40	SHREECEM		
13	CUMMINSIND	27	KOTAKBANK	41	TATAELXSI		

Heatmap for the correlation matrix

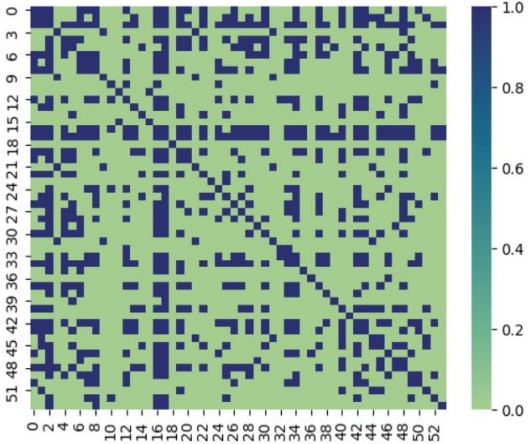


Adjacency matrix for different thresholds

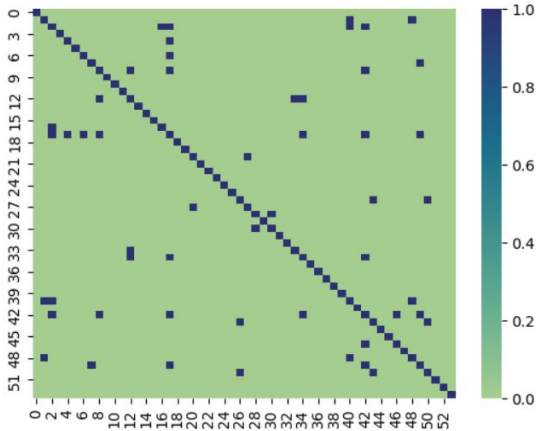
Threshold = 0.25



Threshold = 0.3

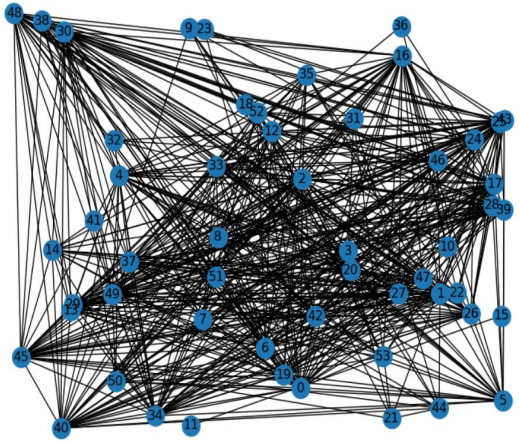


Threshold = 0.5



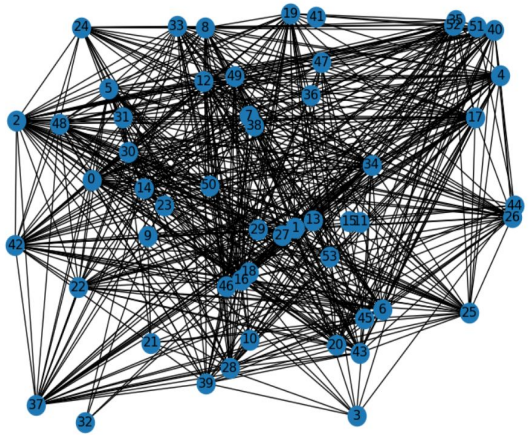
Graph for different Thresholds

Threshold = 0.25



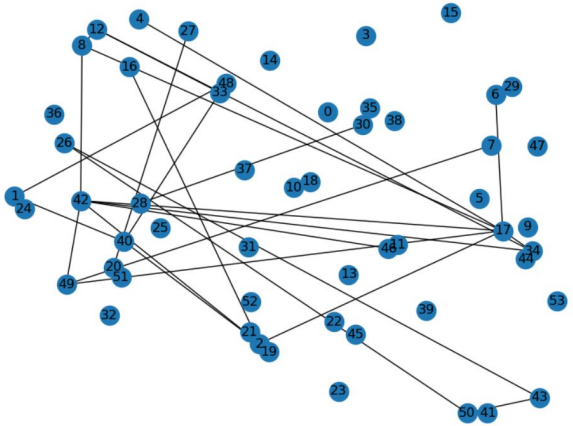
Graph with 54 nodes and 571 edges

Threshold = 0.3



Graph with 54 nodes and 359 edges

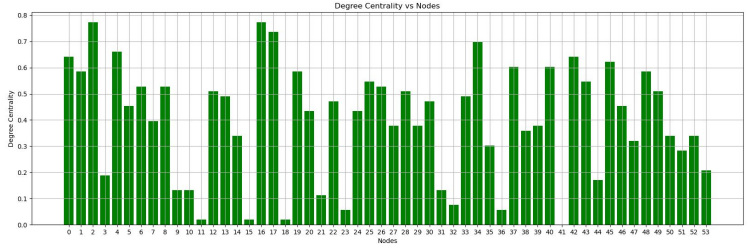
Threshold = 0.5



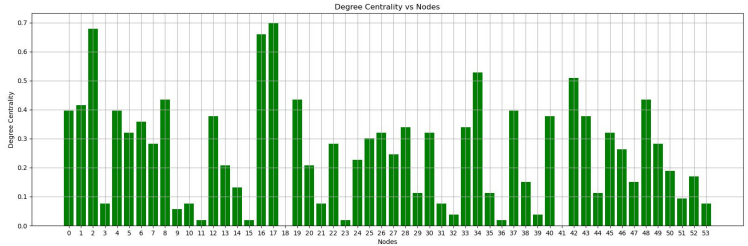
Graph with 54 nodes and 26 edges

Degree Centrality vs Nodes

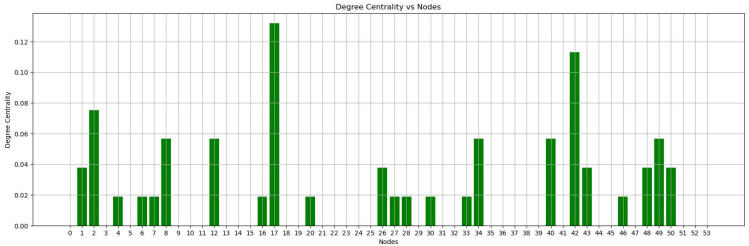
Threshold = 0.25



Threshold = 0.3

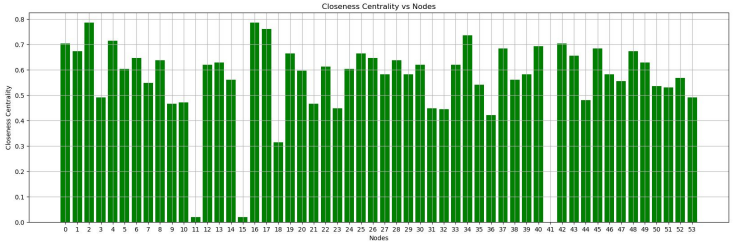


Threshold = 0.5

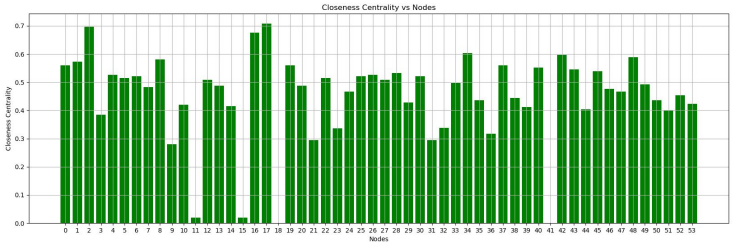


Closeness Centrality vs Nodes

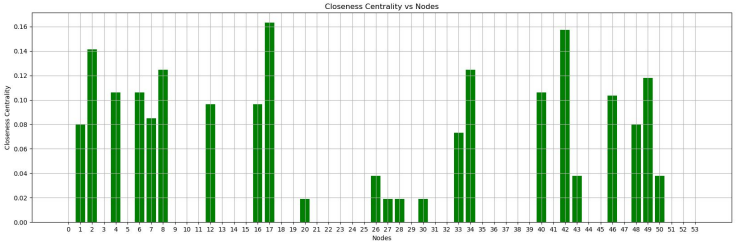
Threshold = 0.25



Threshold = 0.3

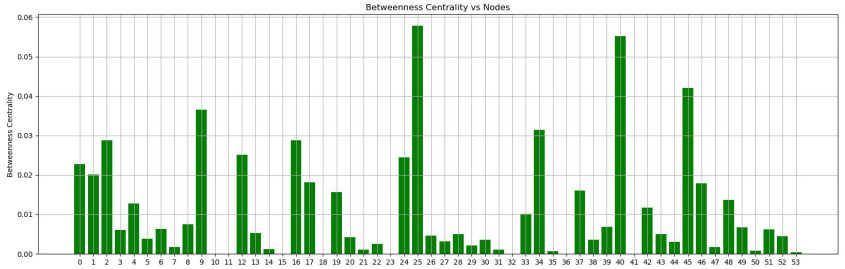


Threshold = 0.5

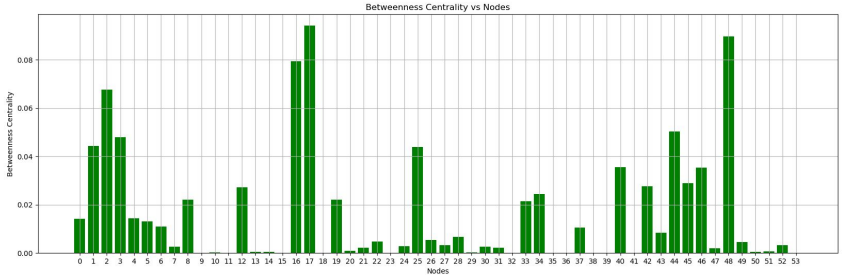


Betweenness Centrality VS Nodes

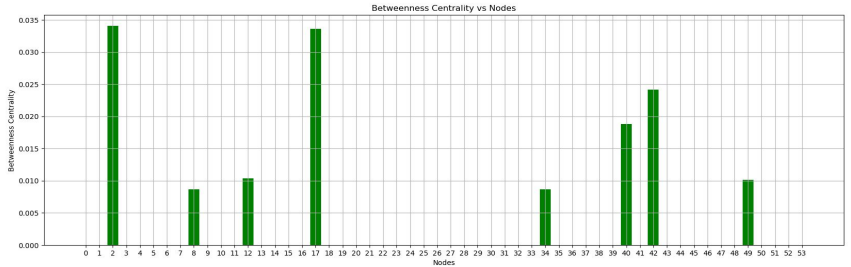
Threshold = 0.25



Threshold = 0.3

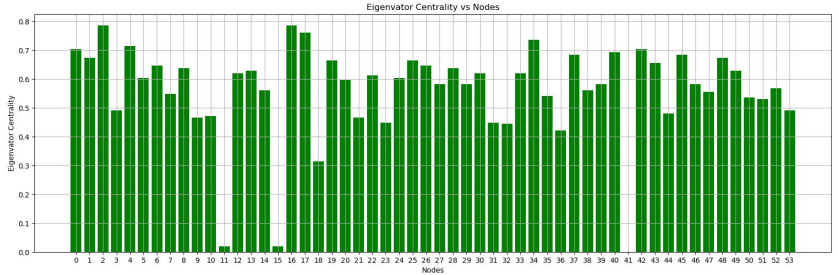


Threshold = 0.5

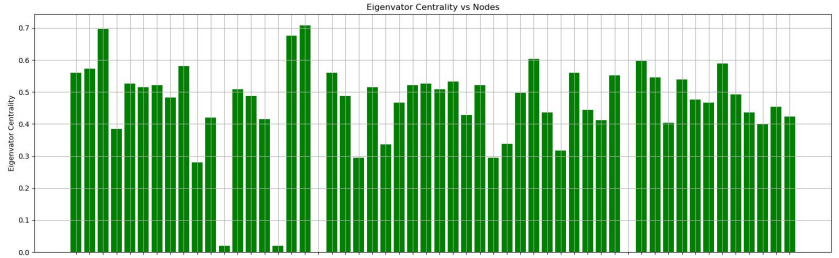


Eigenvector Centrality VS Nodes

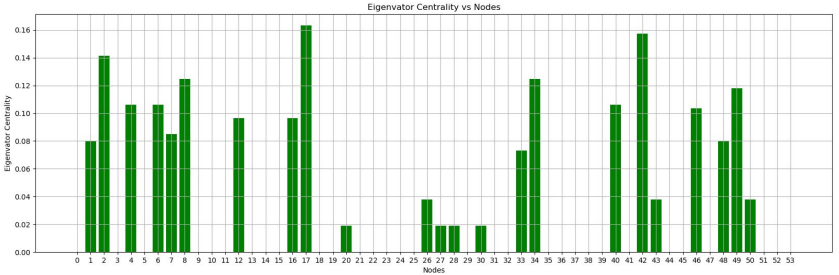
Threshold = 0.25



Threshold = 0.3



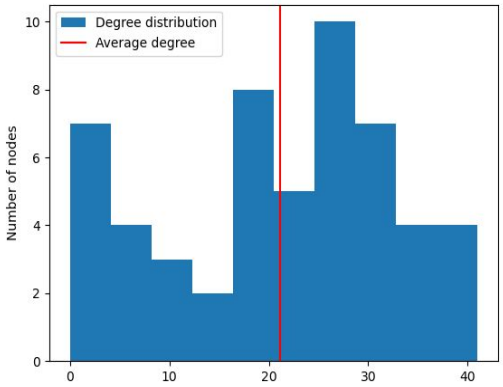
Threshold = 0.5



Degree Distribution and Edge Density for different Thresholds

Average degree is total degree/total number of nodes
Edge density is Total number of edges/ Total possible edges

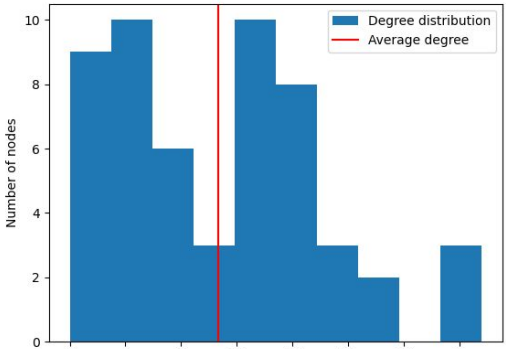
Threshold = 0.25



Average degree 21.14814814814815

Edge density 0.39902166317260657

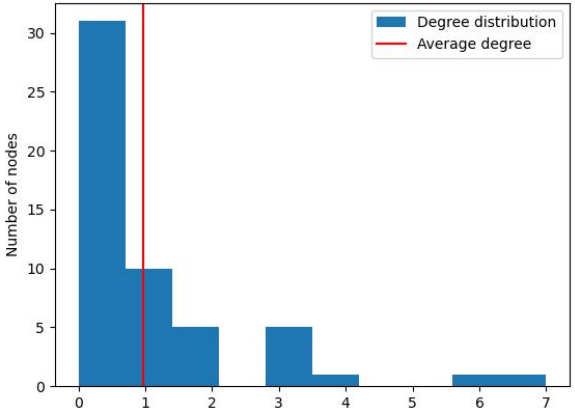
Threshold = 0.3



Average degree 13.296296296296296

Edge density 0.2508735150244584

Threshold = 0.5

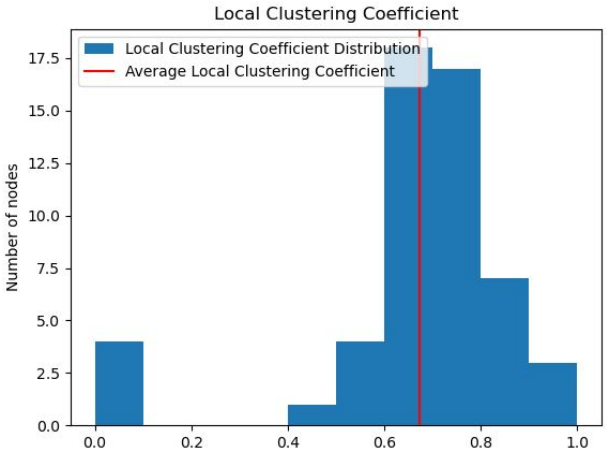


Average degree = 0.9629629629629629

Edge density = 0.01816911250873515

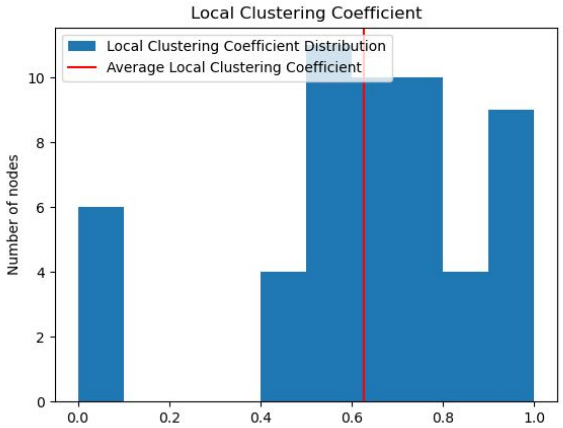
Local Clustering Coefficient for Different Thresholds

Threshold = 0.25



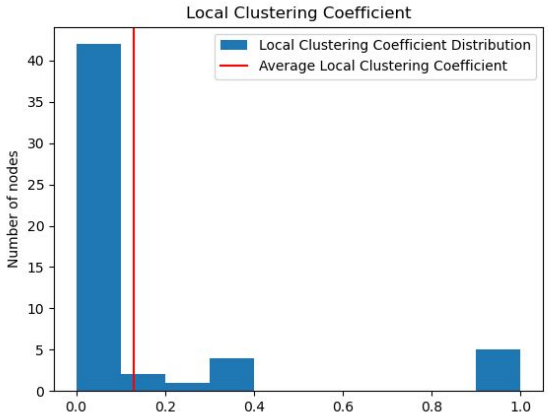
Local Clustering Coefficient
0.6734550455977044

Threshold = 0.3



Local Clustering Coefficient
0.6268233169595075

Threshold = 0.5



Local Clustering Coefficient
0.12883597883597883

Conclusion

- Our code demonstrates the application of network science techniques to analyze financial market data and gain valuable insights into the stock market dynamics by constructing a network based on the correlations, and analyzing the network using various centrality measures.
- The centrality measures help identify influential stocks and their roles within the network.
- Federal Bank gives the highest Degree Centrality, Closeness Centrality and Betweenness Centrality measures for the threshold of 0.5.
- ITC gives the highest Betweenness Centrality measures for the threshold of 0.25.
- Further research and analysis could involve exploring dynamic network properties, community detection, and predictive modeling based on network structure and dynamics.