BACHELOR'S DEGREE PROJECT IN MATHEMATICS

Exploring the Impact of Centrality Measures on Stock Market Performance in Stockholm Market: A Comparative Study

by

Tarek Hasna

MAA043 — Bachelor's Degree Project in Mathematics

DIVISION OF MATHEMATICS AND PHYSICS

MÄLARDALEN UNIVERSITY

SE-721 23 VÄSTERÅS, SWEDEN

Mälardalen University

MAA043 — Bachelor's Degree Project in Mathematics

Date of presentation:

31st May 2023

Project name:

Exploring the Impact of Centrality Measures on Stock Market Performance: A Comparative Study

Author:

Tarek Hasna

Version:

30th June 2023

Supervisor:

Christopher Engström

Examiner:

Linus Calsson

Comprising:

15 ECTS credits

1.1 Abstract

1.1.1 Introduction,

Centrality measures in network analysis have become a popular measurement tool for identifying coherent nodes within a network. In the context of stock markets, the centrality measure helps to identify key performing elements and strengths for specific stocks and determine their impact on disrupting market value and performance. Multiple studies presented practical implementations of centrality measures for determining trends and performance of a particular market. However, fewer studies applied centrality measures to predict trends in the stock market.

1.1.2 Study aim,

This study explores the effectiveness of different centrality measures in predicting trends in the stock market and how the prediction impacts market performance.

1.1.3 literature review,

Studies have shown that centrality measures are applied across many fields [9] [1]. In this study I have reviewed a handful of papers suggesting centrality measures to be effective in predicting stock market performance. The reviewed literature explains how stocks with higher Betweenness Centrality tend to outperform those with lower and that stocks with a higher degree of centrality and eigenvector centrality are found to have a positive relationship with future stock performance. Further research based on the efficient market hypothesis was also reviewed. It shows that analysts can utilise centrality measures to identify influential stock variables [9] [7]. Despite that, studies applied centrality measures to predict trends in the stock market, which we consider a research gap that this study intends to fill.

1.1.4 Methodology,

The study is conducted on secondary data collected from Yahoo on the highest thirty-performing stocks of the Stockholm stock market and based on real-time stock values. Data are analysed using the MATLAB tool. Stock variables were measured based on price and time factors. The measurement is repeated multiple times using six centrality measure types to determine which centrality measure is more effective in determining the actual stock variables. This is found by comparing the results to OMX Stockholm 30 index to determine which centrality measure has the closest correlation coefficient to the OMX Stockholm index.

1.1.5 Conclusion,

All six types of centrality measures have been fundamentally effective in measuring the stock market trend. However, Katz is found to have the closest correlation coefficient to OMX Stockholm.

Keywords: Centrality measures, Degree centrality, Eigenvector centrality, Betweenness centrality, Closeness centrality.

1.2 Acknowledgements

I would like to take a moment to express my heartfelt appreciation to all the individuals who have contributed, in one way or another, to the completion of this study.

First and foremost, I am deeply grateful to my supervisor, Dr. Christopher Engström, for his invaluable guidance, unwavering patience, and insightful feedback. His extensive expertise has been instrumental in shaping this study.

I extend my sincere thanks to my examiner, Dr. Linus Carlsson, for dedicating time and effort to thoroughly read and examine this study. Your commitment ensures that I receive the recognition and evaluation I truly deserve.

I am profoundly touched by the kindness and support extended by my family members and friends. My brothers, Mohammad and Omar, deserve a special mention for their guidance, observations, and assistance in completing this study. Their contributions have been invaluable.

I would also like to express my deepest gratitude to my wife, Khadija, who has been my rock throughout this journey. Her unwavering support and understanding have greatly helped in relieving my stress during the completion of this study.

To all those who have contributed, regardless of the scale of their involvement, I want to acknowledge your significant impact on this study.

Thank you all!

Table of Contents

1.3 Objective and Task

Which centrality measure is the most effective in predicting stock market trends?

1.4 Introduction

Centrality measures are essential in network analysis and have gained popularity in analysing stock markets. This study aims to explore the impact of different centrality measures on stock market performance by comparing their effectiveness in predicting trends, investigating the relationship between centrality and market performance, and identifying potential market risks and opportunities. Data is collected for 73 stocks traded on Nasdaq Stockholm, and their pairwise correlations are calculated to create a graph representation of the network. The results show that Degree Centrality, Betweenness Centrality, Closeness Centrality, Eigenvector Centrality, Katz Centrality, and PageRank Centrality provide valuable insights into the relationships between stocks and can inform investment decisions.

1.5 Literature Review

Centrality measures have become essential for identifying the most significant nodes in networks, including those in the stock market. Degree centrality, Betweenness Centrality, Closeness Centrality, and Eigenvector Centrality are some forms of centrality measurements that can be used to identify significant stocks in a network. Degree centrality measures the number of links a node has to other nodes in the network and can be used to identify highly connected stocks. Betweenness Centrality measures how often a node falls on the shortest path between two other nodes and can be used to identify stocks that serve as a bridge between different sectors or industries [10] . Closeness centrality measures the distance between a node and all other nodes in the network and can be used to identify stocks well connected to other stocks. Eigenvector centrality measures the degree to which a node is connected to other influential nodes in the network and can be used to identify stocks linked to other essential stocks[11].

Although scholarly reviews have yet to be sufficient in covering the practicality of applying centrality measures in predicting stock market trends, a handful of studies have suggested that centrality measures can be an effective method in predicting stock market performance [18] [9]. For example, researchers have found that stocks with higher Betweenness Centrality tend to outperform those with lower. Similarly, stocks with a higher degree of centrality and eigenvector centrality have also been found to have a positive relationship with future stock performance.

The theoretical framework for this study is based on the efficient market hypothesis, which suggests that the stock market is efficient and reflects all available information. However, certain stocks may disproportionately impact market performance, leading to anomalies that investors can exploit. Studies show that centrality measures are used to identify these influential stocks and can therefore be used to test the efficiency of the market [9] [18]. By identifying stocks that are likely to perform well in the future based on their centrality, investors may achieve higher returns than the market as a whole.

Further to the literature on the usability of centrality measures, scholars proposed multiple definitions depending on the field to which the theory is applied [12] [4] [3]. However, all these definitions intersect in a similar description of centrality measures that presents a strong network matrix connecting two or more nodes. The connection between two nodes is presented in edge lines. The significance and closeness of these lines in the path between the nodes represent the strength of this network connection.

Although the literature agreed on a general description of this phenomenon, scholarly definitions have been heterogeneous. According to Newman 2008, centrality measures are defined as quantitative measures used to determine the importance or centrality of a node within a network [12]. In contrast, Krackhardt et al. (2006) described centrality measures as the calculation and evaluation of different measures of centrality in a network [3]. In the first definition, centrality measures are used to determine the importance or prominence of nodes within a network. The second description focuses on the general concept of assessing the significance of nodes within a network. It highlights the specific focus on determining the reliability and performance of centrality measures when dealing with errors and incomplete information in network data. This is because the centrality measure is a broad concept and can be used across disciplines such as Mathematics, Social Networks, and economics. For instance, in a quantitative study to explain network mathematics uses centrality measures in a mathematical technique to explain two main components: the quantitative analysis of network data and the mathematical modeling of networked systems. Newman's research proposes a specific definition of the phenomena based on a given mathematical formula represented by the adjacency matrix [12].

Centrality measures has found application in many fields and studies and has created an arena of socio-science and mathematical measures rather than just a notion. This lack of accurate description of the concept did not prevent scholars from defining the four methods on which centrality measure relies: Degree centrality, Closeness centrality, Betweenness centrality, and Eigenvector centrality .

Types of Centrality Measures & Definition
• General definition by Newman (2008) describes centrality measures as a network made up of points in the form of nodes or vertices connected by lines, referred to as a graph in the mathematics literature. This network is represented by an adjacency matrix and mathematically represents relationships between nodes [12]. • Degree centrality, defined by Freeman (1979), is the number of ties incident to a node, represented by the total of each row in the network's adjacency matrix. More formally, it is the number of paths of length one that emanate from a node [22]. • Krackhardt (1990) extends the centrality measure definition to identify the most central nodes in a network by measuring betweenness centrality. Betweenness centrality measures the extent to which a node lies on the shortest paths between other nodes in a network. It is commonly used to find nodes that connect distinct areas of a graph [3]. • Eigenvector centrality is used to identify and measure the magnitude of network connections. It helps understand the strength of connections between two nodes. Cheung et al. (2020) interpret eigenvector centrality in terms of direct and indirect connections between nodes, assuming a network is strongly connected if there is a direct path between any two distinct nodes [6].
Application of Centrality Measures The literature shows that centrality measures are widely used in various fields such as social science, mathematics, and economics [6] [12]. These measures are employed as analytical tools to study stock market performance by analyzing the behavior of the stock market over a specific time period. The four centrality measures mentioned earlier (degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality) are used to predict stock market trends. The analysis is evaluated and compared to determine the most effective method for predicting stock market trends.

1.6 Methods

1.5.1 Introduction,

This chapter discusses the methodology used to conduct the analysis of this research, based on quantitive research technique. The study is a comparative analysis; it compares the stock market trends and behaviour in Stockholm to the different types of centrality measuring outlined in the previous chapters. The research is based on secondary data collected from reliable webpage and relevant literatures that have discussed similar topic.

1.5.2 Research technique,

The research is a quantitative analysis that measures pricing over time. Precisely, it measures the closing price of all shares in the stock market during the specified period. Closing pricing is used to measure the daily return for all stock shares, through calculating the difference between the "In price" of the last day and the "Closing price" of the previous day. The daily return of all stock is then applied to work out the correlation matrix that presents the correlation coefficients between a group of variables.

1.5.3 Data collection,

This research is conducted using secondary data, collected from Yahoo webpage that contains stock share pricing trends, including "in price" and "closing price" for shares for a defined period.

1.5.4 Tools,

Measurements and analysis are delivered using the MATLAB tool, a numeric computing platform used to generate a script of market data from the secondary data. Secondary data is imported from the Yahoo webpage into MATLAB drive. Command codes are then generated to develop/import a script of market data from MATLAB drive. Measurements are based on pre-defined formulas generated in MATLAB and extracted from different types of centrality measures listed hereunder;

1.5.4 - 1 Degree Centrality:

Degree centrality is a measure of the importance of a node in a network, based on the number of connections (edges) it has with other nodes. It is defined as the total number of edges that are incident on a node [19].
The degree centrality of node i is given by:

$$\text{Degree Centrality of Node } i = \text{Number of edges connected to node } i.$$

1.5.4 - 2 Eigenvector Centrality:

Eigenvector centrality is a commonly used method for determining the significance of nodes in a graph. It is computed by finding the eigenvector of the graph's adjacency matrix, with the entry in the eigenvector corresponding to a node representing that node's eigenvector centrality. If an eigenvalue λ of a matrix A is positive, the standard definition of eigenvectors $\mathbf{v}$ states that when multiplied by matrix A, eigenvectors belonging to this eigenvalue only change in scale, but their direction remains unchanged (i.e., $A\mathbf{v} = \lambda\mathbf{v}$). Usually, when a vector is multiplied by a matrix, it changes direction, but eigenvectors retain their direction when multiplied by the matrix if the eigenvalue is positive. In other words, when an eigenvector is multiplied by a matrix, the resulting vector is simply a scaled version of the original eigenvector and points in the same direction [2] [20].

$$\text{centrality}(v) = \frac{1}{\lambda} \sum_{j=1}^{n} a_{j,v} \text{centrality}(j), \tag{1}$$

where $a_{j,v}$ is the adjacency matrix element between nodes j and v, and λ is a normalisation factor [14].

Example 1: Calculation of Eigenvector Centrality

Consider a direct graph with 3 nodes, Node 1, Node 2 and Node 3, with an eigenvalue of 2 (i.e. $\lambda = 2$). We have the following connections:

Node 1 is connected to Node 2 and Node 3.

Node 3 is connected to Node 2.

We want to calculate the Eigenvector centrality for each node in the graph using formula 1: We create the following adjacency matrix:

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

We initialize the centrality vector with equal values:

$$v_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Iteration 1:

For node 1:

$$\begin{aligned} \text{centrality}(node1) &= \frac{1}{\lambda}(0 \cdot \text{centrality}(node1) + 1 \cdot \text{centrality}(node2) + 1 \cdot \text{centrality}(node3)) \\ &= \frac{1}{2}(1 + 1) = 1. \end{aligned}$$

For node 2:

$$\begin{aligned} \text{centrality}(node2) &= \frac{1}{\lambda}(0 \cdot \text{centrality}(node1) + 0 \cdot \text{centrality}(node2) + 0 \cdot \text{centrality}(node3)) \\ &= 0. \end{aligned}$$

For node 3:

$$\begin{aligned}\text{centrality}(node3) &= \frac{1}{\lambda}(0 \cdot \text{centrality}(node1) + 1 \cdot \text{centrality}(node2) + 0 \cdot \text{centrality}(node3)) \\ &= \frac{1}{2}(1) = \frac{1}{2}.\end{aligned}$$

After the first iteration, the resulting centrality vector is:

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ \frac{1}{2} \end{bmatrix}.$$

Iteration 2:

For node 1:

$$\begin{aligned}\text{centrality}(node1) &= \frac{1}{\lambda}(0 \cdot \text{centrality}(node1) + 1 \cdot \text{centrality}(node2) + 1 \cdot \text{centrality}(node3)) \\ &= \frac{1}{2}(0 + \frac{1}{2}) = \frac{1}{4}.\end{aligned}$$

For node 2:

$$\begin{aligned}\text{centrality}(node2) &= \frac{1}{\lambda}(0 \cdot \text{centrality}(node1) + 0 \cdot \text{centrality}(node2) + 0 \cdot \text{centrality}(node3)) \\ &= 0.\end{aligned}$$

For node 3:

$$\begin{aligned}\text{centrality}(node3) &= \frac{1}{\lambda}(0 \cdot \text{centrality}(node1) + 1 \cdot \text{centrality}(node2) + 0 \cdot \text{centrality}(node3)) \\ &= \frac{1}{2}(0) = 0.\end{aligned}$$

After the second iteration, the resulting centrality vector is:

$$v_2 = \begin{bmatrix} \frac{1}{4} \\ 0 \\ 0 \end{bmatrix},$$

From the previous example we can conclude that:

After the first iteration, the centrality values for the nodes are: Node 1 = 1, Node 2 = 0, and Node 3 = $\frac{1}{2}$. Node 1 has the highest centrality value, indicating that it is the most central node in the graph at this stage.

After the second iteration, the centrality values are: Node 1 = $\frac{1}{4}$, Node 2 = 0, and Node 3 = 0. Node 1 still has the highest centrality value, although it has decreased. Node 2 and Node 3 both have centrality values of 0, suggesting that they are less central in the graph.

If we were to proceed with further iterations, the centrality values would continue to change, potentially converging to a stable solution. However, in this specific example, we have stopped at the second iteration, and the resulting centrality vector is

$$v_2 = \begin{bmatrix} \frac{1}{4} \\ 0 \\ 0 \end{bmatrix}.$$

1.5.4 - 3 Katz Centrality:

Katz centrality is a node centrality metric that takes into account a node's immediate and distant neighbors. It is described as the total of the weighted paths of length k connecting a node to every other node in the network, where k is a parameter that determines the degree of influence of indirect paths on the centrality measure. In social network research, recommendation engines, and other applications where the importance of indirect connections between nodes is significant, Katz centrality is frequently utilized. It has been demonstrated to be very helpful for locating influential nodes in big, complicated networks [24].

$$C_{Katz}(v_i) = \alpha \sum_{j=1}^{n} a_{j,i} C_{Katz}(v_j) + \beta, \tag{2}$$

C_{Katz}	is the Katz Centrality of a node v_i.
α	is the damping factor.
β	is a constant term.
a	represents the adjacency matrix.
n	the total number of nodes in the network.

[13].

Example 2: Calculating Katz Centrality

Consider a direct graph with 3 nodes, Node 1, Node 2 and Node 3. We have the following connections:

Node 1 is connected to Node 2.

Node 3 is connected to Node 2.

We will use the following parameter values to calculate Katz centrality using formula (2):

Damping factor, $\alpha = 0.85$

Constant term, $\beta = 1$

Number of iterations = 2

We create the adjacency matrix, A: The adjacency matrix will be:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Initialize the Katz Centrality scores, C_{Katz}, for each node: Start with an initial value of 1 for all nodes.

$$C_{\text{Katz}} = [1, 1, 1].$$

Iterate the Katz Centrality calculation until convergence: Apply the formula to update the Katz Centrality scores for each node until convergence. Let's go through one iteration of the calculation:

For node 1:

$$C_{Katz}(v_1) = \alpha \sum_{j=1}^{n} A_{1,j} C_{Katz}(v_j) + \beta = 0.85 \cdot (A_{1,1} C_{Katz}(v_1) + A_{1,2} C_{Katz}(v_2) + (A_{1,3} C_{Katz}(v_3)) + 1$$
$$= 0.85 \cdot (0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1) + 1 = 1.85.$$

For Node 2:

$$C_{Katz}(v_2) = \alpha \sum_{j=1}^{n} A_{2,j} C_{Katz}(v_j) + \beta = 0.85 \cdot (A_{2,1} C_{Katz}(v_1) + A_{2,2} C_{Katz}(v_2) + (A_{2,3} C_{Katz}(v_3)) + 1$$
$$= 0.85 \cdot (0 \cdot 1 + 0 \cdot 1 + 0 \cdot 1) + 1 = 1.$$

For Node 3:

$$C_{Katz}(v_3) = \alpha \sum_{j=1}^{n} A_{3,j} C_{Katz}(v_j) + \beta = 0.85 \cdot (A_{3,1} C_{Katz}(v_1) + A_{3,2} C_{Katz}(v_2) + (A_{3,3} C_{Katz}(v_3)) + 1$$
$$= 0.85 \cdot (0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1) + 1 = 1.85.$$

After one iteration, the updated Katz Centrality scores are:

$$C_{\text{Katz}} = [1.85, 1, 1.85].$$

We do the same steps with updated Katz centrality.

Iteration 2:

For Node 1:

$$C_{Katz}(v_1) = \alpha \sum_{j=1}^{n} A_{1,j} C_{Katz}(v_j) + \beta = 0.85 \cdot (A_{1,1} C_{Katz}(v_1) + A_{1,2} C_{Katz}(v_2) + (A_{1,3} C_{Katz}(v_3)) + 1$$

$$= 0.85 \cdot (0 \cdot 1.85 + 1 \cdot 1 + 0 \cdot 1.85) + 1 = 1.85.$$

For Node 2:

$$C_{Katz}(v_2) = \alpha \sum_{j=1}^{n} A_{2,j} C_{Katz}(v_j) + \beta = 0.85 \cdot (A_{2,1} C_{Katz}(v_1) + A_{2,2} C_{Katz}(v_2) + (A_{2,3} C_{Katz}(v_3)) + 1$$

$$= 0.85\c{(}0 \cdot 1.85 + 0 \cdot 1 + 0 \cdot 1.85) + 1 = 1.$$

For Node 3:

$$C_{Katz}(v_3) = \alpha \sum_{j=1}^{n} A_{3,j} C_{Katz}(v_j) + \beta = 0.85 \cdot (A_{3,1} C_{Katz}(v_1) + A_{3,2} C_{Katz}(v_2) + (A_{3,3} C_{Katz}(v_3)) + 1$$

$$= 0.85\c{(}0 \cdot 1.85 + 1 \cdot 1 + 0 \cdot 1.85) + 1 = 1.85.$$

After iteration 2, the updated Katz Centrality scores are:

$$C_{Katz} = [1.85, 1, 1.85].$$

From the previous example we can conclude the following:

After the first iteration, the Katz centrality scores for the nodes were updated as follows: Node 1: 1.85, Node 2: 1, Node 3: 1.85. These updated scores reflect the influence and importance of each node within the graph based on their connections and the Katz centrality algorithm.

In the second iteration, the Katz centrality scores were recalculated using the updated scores from the first iteration. The resulting scores were identical to the scores obtained after the first iteration: Node 1: 1.85, Node 2: 1, Node 3: 1.85. This indicates that the Katz centrality algorithm has reached convergence, and further iterations would not change the scores significantly.

1.5.4 - 4 Betweenness Centrality:

A node's relevance in a network is gauged by its capacity to act as a bridge along the shortest routes connecting pairs of other nodes, a concept known as Betweenness Centrality. The number of shortest paths that pass through a node to connect all pairs of nodes in the network is what is known as a node's Betweenness Centrality. High Betweenness Centrality nodes are seen to be more significant in the network since their removal could obstruct information flow and communication between other nodes. In social network research, transportation planning, and other fields where it is crucial to understand how information or resources move across a network, Betweenness Centrality is frequently utilized [24].

$$\text{Betweenness Centrality}(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

σ_{st}	is the total number of shortest paths from node s to node t.
$\sigma_{st}(v)$	is the number of shortest paths from node s to node t that pass through node v.

[5].

1.5.4 - 5 Closeness centrality:

Based on a node's average distance from all other nodes in the network, a measure of a node's significance in a network is called closeness centrality[25]. The reciprocal of the sum of a node's shortest path distances to every other node in the network is the closeness centrality of that node. Because they can rapidly and effectively reach other nodes in the network, nodes with high closeness centrality are thought to be more significant. In transportation design, it is crucial to locate central places that can give easy access to every other location in the network. This concept is known as closeness centrality[24].

$$\text{Closeness Centrality}(i) = \frac{N-1}{\sum_{j=1}^{N} d(i,j)},$$

N	is the total number of nodes in the graph.
$d(i,j)$	is the shortest path distance between node i and node j.
$\sum_{j=1}^{N} d(i,j)$	denotes the sum of the shortest distances between node i and all other nodes j in the network.

[16].

Since the graph in this study is not strongly connected, to handle that, we introduce two additional terms, A_i and $\sum_{j \in A_i} d(i,j)$.

A_i represents the number of reachable nodes from node "i," while $\sum_{j \in A_i} d(i,j)$ represents the sum of distances from node "i" to all reachable nodes.

Now, let's modify the numerator and denominator of the first formula:

For the numerator, we replace $N - 1$ with A_i. This change accounts for the fact that in a not strongly connected graph, not all nodes may be reachable from node "i." So, instead of considering the total number of nodes, we focus only on the reachable nodes.

For the denominator, we replace $\sum_{j=1}^{N} d(i, j)$ with $\sum_{j \in A_i} d(i, j)$. This adjustment ensures that we only consider the sum of distances to reachable nodes from node "i," excluding any unreachable nodes.

By making these modifications, we arrive at the second formula for closeness centrality:

$$\text{Closeness Centrality}(i) = \left(\frac{A_i}{N-1}\right)^2 \cdot \frac{1}{\sum_{j \in A_i} d(i, j)}.$$

[21].

In this modified formula, $\left(\frac{A_i}{N-1}\right)^2$ represents the ratio of reachable nodes from node "i" out of all possible nodes except for node "i". Squaring this ratio emphasizes the importance of having a higher number of reachable nodes.

The term $\frac{1}{\sum_{j \in A_i} d(i,j)}$ in the formula adjusts the centrality measure based on the sum of distances from node "i" to reachable nodes. It implies that a shorter average distance to reachable nodes results in a higher closeness centrality.

Therefore, the second formula incorporates the concepts of reachable nodes and distances to adjust the closeness centrality calculation for graphs that are not strongly connected.

1.5.4 - 6 PageRank:

The mathmatical PageRank algorithm is a formula which is used to evaluate the significance of nodes in a network or graph. Each node in an undirected graph is set a number from the algorithm based on the quantity and quality of the inbound linkages from other nodes in the network. The algorithm takes into account both the number of incoming links and the importance of the linking nodes. Nodes that receive many high-ranking inbound links are considered more important and will have higher PageRank values. By iteratively evaluating and updating these values, the algorithm ultimately provides a measure of the relative importance of nodes within the network. This methodology has widespread applications across various domains, including search engines, chemistry, social network analysis, molecular biology, and more [26] [8].

PageRank $\overrightarrow{R}$ for the vertices in a graph $G := (V; E)$ is defined as the (right) eigenvector with eigenvalue one to the matrix M:

$$\mathrm{M} = c(a + \overrightarrow{g}\,\overrightarrow{w}^{\top})^{\top} + (1-c)\overrightarrow{w}\,\overrightarrow{e}^{\top} \quad ,$$

a	is the adjacency matrix weighted such that the sum over every nonzero row is equal to one.
$\overrightarrow{g}$	is a vector with zeros for vertices with outgoing edges and 1 for all vertices with no outgoing edges.
$\overrightarrow{w}$	is a non-negative vector with norm $\lVert\overrightarrow{w}\rVert_1 = 1$.
$\overrightarrow{e}$	is a one-vector and $0 < c < 1$ is a scalar.
$\top$	Denotes the transpose of a vector or matrix.

Where the term $(A \overrightarrow{+} g \overrightarrow{w} \top)$ denotes the adjusted adjacency matrix. It incorporates the connectivity information (A) and includes outgoing edges for dangling vertices $\overrightarrow{(} g \overrightarrow{w} \top)$, ensuring that the resulting matrix is stochastic. The term $(1 - c)\overrightarrow{w} e \top$ represents the teleportation vector. It introduces the random jump factor and helps make the matrix positive, irreducible and stochastic [8].

1.7 Analysis:

Data is imported from Yahoo Finance to the MATLAB platform. The data is daily historical data of the stock markets in Stockholm and was used in this study from April 23-2022 till April 23-2023. The data comprises information on seventy-three (73) stocks and presents the characteristics per each share in a matrix consisting of seven variables and 256 observations (256×7), where 256 represents the number of trading days during the specified period, and 7 represents the Date, Open, High, Low, Close, Adjustment, and Volume price in Swedish currency (Swedish Krona SEK).

However, since the research intends to compare the daily return, only the close column information is needed. Hence, closing price data of all stocks for a specified period is imported for this observation.

Market Data:	Is a 256×73 matrix that contains the closing price of all shares during the specified period.
S:	Refers to the number of stocks imported to conduct this study.
ts (time series):	It is the time-series specified for this study
Symbols:	Refers to the stock symbols according to their importing order.

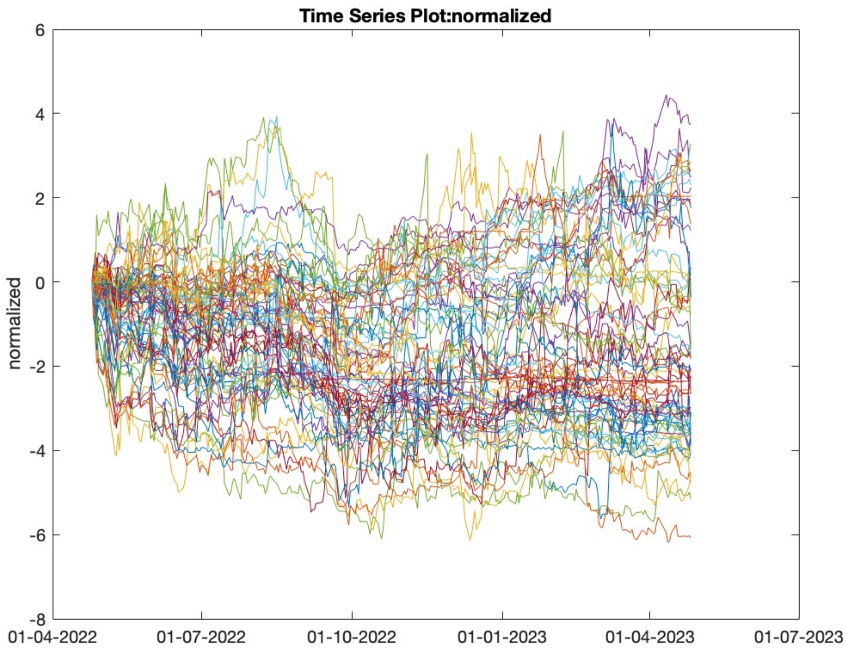

Figure 1: Normalized Prices for 73 Stocks Over Time

After importing the data, the first step is calculating the daily returns for all stocks. The daily return for each stock is determined by subtracting the natural logarithm of the closing price of the previous day from the natural logarithm of the closing price today (i.e., "ln closing price today" - "ln closing price the previous day").
For example let's assume we have the closing prices of a stock for two consecutive days, denoted as $P(n)$ and $P(n+t)$. To calculate the daily return r:

$P(n+t) = P(n) \cdot e^{rt}$,

where $t = 1$,

then: $e^r = \frac{P(n+1)}{P(n)}$,

$ln\ e^r = ln\ (\frac{P(n+1)}{P(n)})$,

$r = ln\ (P(n+1)) - ln\ (P(n))$.

Finding the daily return is crucial for analysing the market's fluctuations since the size of the daily return determines the percentage of price volatility. A high daily return indicates a more significant volatility percentage in price percentage. Similarly, a positive daily return indicates an increase in the price, and a negative return indicates a decrease in the price. Calculating the daily return of all stock is also necessary for the correlation matrix, which presents the correlation coefficients between the variables.

A correlation matrix summarizes the relationships between variables. The correlation between variables is measured using a correlation coefficient, such as Pearson's r, which is used in this study, the correlation coefficients measure the strength and direction of the relationship between the variables. And, since this study uses (an undirected graph), it will focus only on the strength of the relationship between the variables (stocks).
The Pearson's correlation coefficient (r) between two variables, x and y, can be computed using the following formula:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

where:

x_i	represents individual data points of the variable x.
y_i	represents individual data points of the variable y.
$\bar{x}$	represents the mean (average) of x
$\bar{y}$	represents the mean (average) of y.
n	represents the total number of observations.

[15].

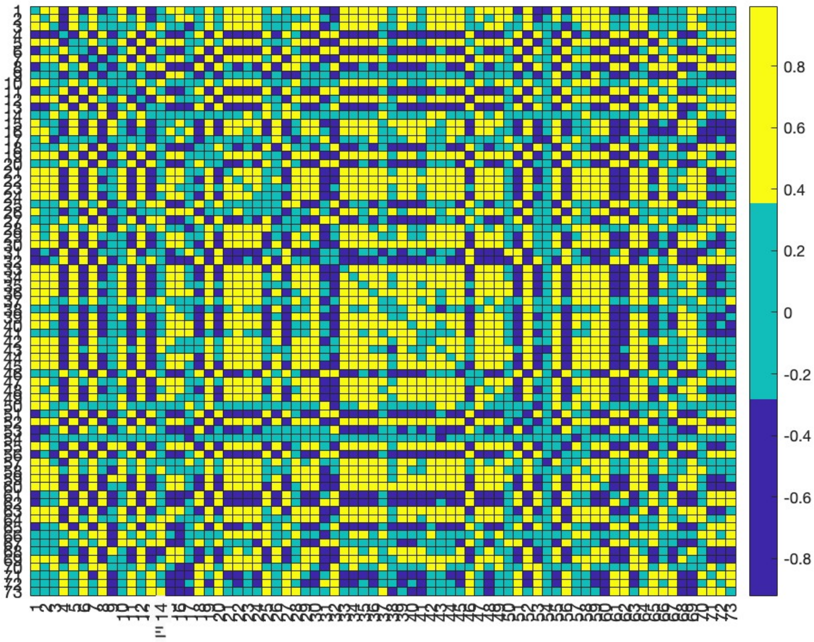

Figure 2: Heat Map of Correlation Matrix For The Market Data

The correlation coefficient is in the range of (-1, +1), +1 correlation coefficient indicates a completely positive correlation. Where a positive correlation coefficient between two variables, x and y, indicates that their relationship is proportional (i.e. when variable x increases, so does variable y). Where -1 correlation coefficient indicates an entirely negative correlation (the two variables have an inverse proportional relationship). This implies that as variable x increases, variable y decreases (in the opposite direction). While the correlation coefficient of zero is indicate no correlation between the two variables[15].

It is crucial to understand that the correlation coefficient, although it can indicate the strength and direction of the relationship between variables, does not establish causation or provide a comprehensive understanding of how the variables interact. In simpler terms, even if there is a perfect correlation coefficient between two variables, it does not necessarily mean that a change in one variable will lead to a change in the other variable. The correlation coefficient alone does not offer specific insights into the precise nature of their relationship and interaction.

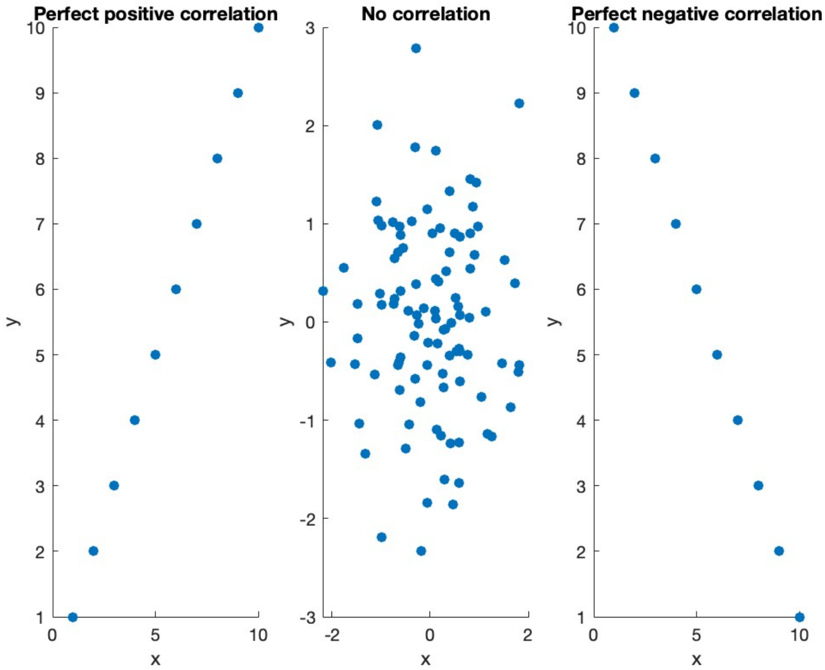

Figure 3: Types of Correlations

We can calculate the adjacency matrix once the correlation matrix has been calculated, which is a square matrix, represents the connections or relationships between nodes in a graph.
Through the following code, we can find the adjacency matrix in MATLAB,

```
adjacency_matrix = (corr_matrix > 0.2)- eye (s);
```

The threshold value was chosen to be 0.2. This means that stocks with relatively weak correlations with a correlation coefficient of greater than 0.2 will be considered correlated in the resulting adjacency matrix. The threshold value of 0.2 represents the smallest value of threshold that ensures enough nodes in the graph are connected to each other. As the threshold increases, the number of stocks considered connected decreases.
Where Threshold is set to 0.2 meaning that any value less than or equal to 0.2 will be set as 0 and any correlation value greater than 0.2 will be set as 1. $Eye(s)$ is the diagonal diameter of the correlation matrix and we subtract it to ensure the nodes don't connect themselves, because in the correlation matrix the diagonal is always equal to 1.

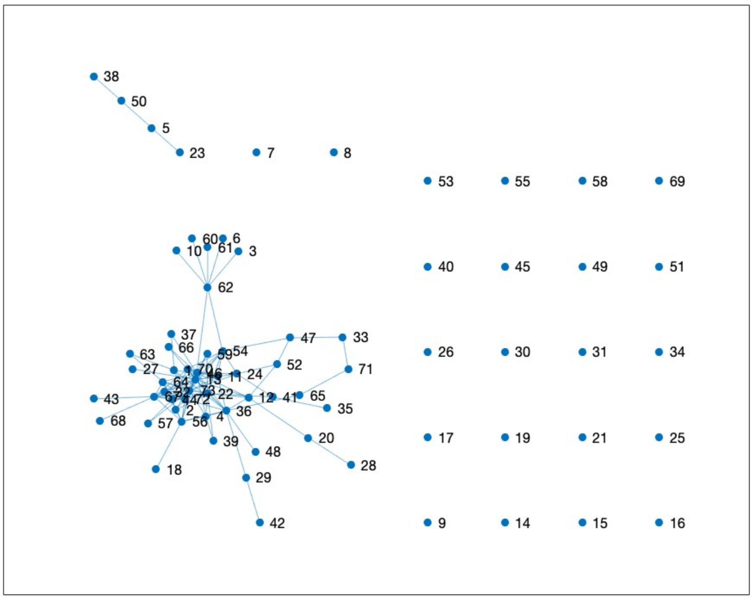

Figure 4: Graph Object of Adjacency Matrix

The graph corresponding to the calculated adjacency matrix can be observed in Figure 4, which represents the relationships among the 73 stocks. In this adjacency matrix, each row corresponds to a specific node in the graph, while an element $a_{i,j} = 1$ in the matrix signifies a link from vertex i to vertex j. The presence of a link between two nodes indicates a correlation value of 1 between the corresponding stocks.

After we get the adjacency matrix, the correlation matrix, and the graph of the adjacency matrix, we can find: Degree, Betweenness, Closeness, Eigenvector, Katz, and PageRank centrality.

Degree centrality:

Can be calculated using the MATLAB tools, which utilizes the centrality function to calculate the degree centrality for each node in the graph. The degree centrality of a node i is defined as:

The normalised degree centrality of node i which used for undirected graph is given by:

$$\text{degree centrality}(i) = \frac{k_i}{N-1} = \frac{\sum_{j \in G} A_{ij}}{N-1},$$

Where, k_i is the degree of node i, A represent the adjacency matrix, N the total number of nodes in the graph, and G is the graph [23].

Betweenness Centrality:

We calculate the Betweenness Centrality of a specific node in the network by calculating the shortest paths between all pairs of nodes in the network that pass through this node divided by the total number of shortest paths between pairs of nodes in the network.

$$\text{Betweenness Centrality}(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

where σ_{st} is the total number of shortest paths from node s to node t, and $\sigma_{st}(v)$ is the number of shortest paths from node s to node t that pass through node v [5].

We can calculate the shortest path distances using algorithms such as Floyd-Warshall or Dijkstra. To calculate Betweenness centrality we used the function "centrality" in MATLAB. Although MATLAB does not clearly disclose the method it employs, the function `centrality(graph, 'betweenness')` computes the betweenness centrality of nodes in the graph object. Betweenness centrality is a measure of the number of times a node acts as a bridge along the shortest path between two other nodes.

Closeness Centrality:

The closeness centrality of a node is calculated by the formula:

$$C_i = \left(\frac{A_i}{N-1}\right)^2 \cdot \frac{1}{C_i},$$

C_i	is the sum of distances from node i to all reachable nodes.
N	is the number of nodes in the graph.
A_i	is the number of reachable nodes.

[21].

To calculate the closeness centrality for a specific node in a graph, we need to calculate the shortest path distance between the node and all other nodes. We can calculate the shortest path distances using an algorithm like Floyd-Warshall or Dijkstra. Alternatively, we can use the MATLAB tools to calculate the closeness centrality, which is a measure of how easy it is for a node to reach other nodes in the graph.

Eigenvector centrality:

We calculate the centrality of node v in a network by the formula:

$$\text{centrality}(v) = \frac{1}{\lambda} \sum_{j=1}^{n} a_{j,v} \text{centrality}(j),$$

where $a_{j,v}$ is the adjacency matrix element between nodes j and v, and λ is a normalization factor [14] [17].

To make the calculation, we do an iterative process where random values are assigned to the degrees of centrality, and then it is done with the equation until it converges with the correct values. There are several ways to implement this process, such as the power method algorithm.

The convergence of the eigenvector centrality algorithm is influenced by several factors. Graph properties play a role, as strongly connected graphs are more likely to yield convergence compared to graphs with disconnected components. The initialization of centrality values and the choice of the normalization factor also impact convergence. The initial values assigned to each node's centrality and the normalization factor λ, which balances local and global connectivity, affect the rate and quality of convergence.

The convergence criterion determines when the algorithm considers the centrality values to have converged. This criterion can be a threshold, where the algorithm stops iterating once the centrality values change less than a predefined threshold, or a maximum number of iterations. The specific criterion chosen depends on the desired level of accuracy and efficiency.

However, it's important to note that eigenvector centrality is an approximation and may not always converge to the "correct" centrality values. Interpretation of centrality values should consider the specific application or analysis being conducted and should be cross-validated with other centrality measures or external knowledge when possible.

By MATLAB we calculate the eigenvector centrality scores for the graph represented by its adjacency matrix using the centrality function.

PageRank centrality:

The PageRank algorithm is a method used to determine the importance of nodes in a network. In the code provided, we use the built-in 'centrality' function in MATLAB's Graph Theory Toolbox to calculate the PageRank centrality of a network represented by a graph G. The following code demonstrates how to calculate PageRank with a damping factor of 0.85 and a total of 100 iterations:

```
damping_factor = 0.85;
num_iterations = 100;
pr = centrality(G, 'pagerank', 'MaxIterations', num_iterations,
'FollowProbability', damping_factor);
```

The damping factor in PageRank determines the probability that a user will continue to follow links in the network, rather than randomly jumping to another page, in stock networks the damping factor It represents the probability of the user following the conections between stocks in the stocks network rather than moving from one stock to a randomly chosen different stock . A damping factor of 0.85 is commonly used in PageRank because it has been found to balance between the importance of high-degree nodes and the exploration of new nodes. However, it is worth noting that for future studies, it would be beneficial to conduct further investigation and examine different damping factor values to assess their impact on the results.

In this study, n this study, a sensitivity analysis was applied by trying different damping factor values (e.g., 0.75, 0.8, 0.85, 0.9, etc.). The purpose was to compare the stocks with the highest centrality across the PageRank measure (using a different damping factors) with the results obtained using other centrality measures such as Degree, Closeness, Betweenness, Katz, and Eigenvector centrality. The sensitivity analysis found that the two stocks with the highest centrality across the PageRank measure, using a damping factor of 0.85, were Prevas B and Karolinska Development B. These two stocks matched exactly with the stocks having the highest centrality according to the Degree , Eigenvector, and Closeness centrality measures. However, when using other damping factor values like 0.7, the stocks with the highest centrality were not similar to those identified by other centrality measures.

Regarding the number of iterations in the PageRank algorithm, it determines how long the algorithm runs before producing a result. In this study, a large number of iterations was chosen, along with a moderately sized damping factor, to mitigate potential issues with algorithm convergence.

Katz centrality:

We start by creating the weighted adjacency matrix with the weighted edges, the weighted matrix used in the Katz centrality algorithm enables a more complex examination of the significance and centrality of graph nodes. Instead of only examining whether a connection exists or not (as in an unweighted matrix), then we take the average of the weighted adjacency matrix and its transpose to create an undirected adjacency matrix. This is done to ensure that the graph is symmetric. We create a graph object based on the undirected adjacency matrix.

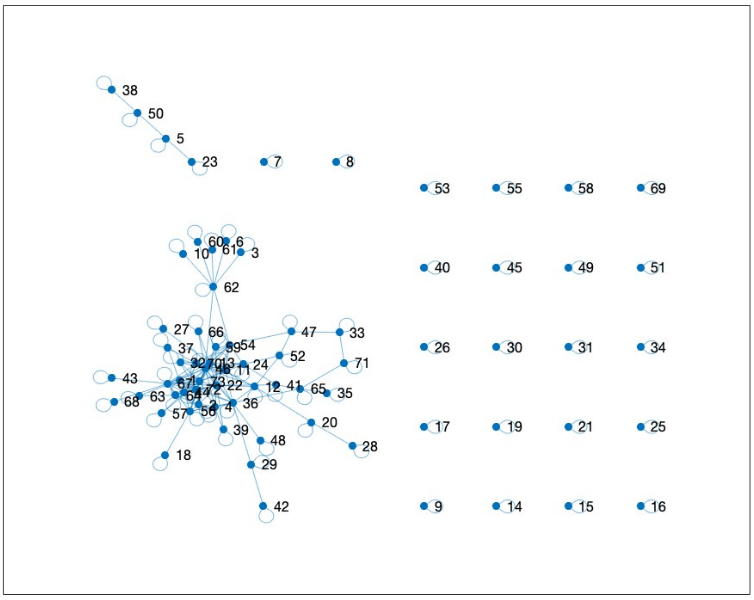

Figure 5: Graph Object of Weighted Adjacency Matrix

We set a value for alpha (0.18), which is a damping factor used in the Katz centrality algorithm. By using MATLAB function we initialize the centrality value for each node to 1.

```
AK = alpha * A * Katz;
for k = 1:100
AK = alpha * A * AK;
Katz = Katz + AK;
end
Katz=Katz+1;
```

By pervious codes we performs 100 iterations of the Katz centrality algorithm to calculate the centrality value of each node.

This code performs 100 iterations of the Katz centrality algorithm to calculate the centrality value of each node. The resulting centrality values are then added by 1 to ensure that all values are positive.

After calculating the centrality of the stocks through all the measures (Degree, Katz, Betweenness, Closeness, PageRank, and Eigenvector), we normalise the centrality measures to visualize and compare the measures in the graph. Then we normalize the results of centrality obtained through the Katz centrality measure, normalization is done by dividing each Katz centrality value by the sum of all Katz centrality values. The normalization process ensures that the sum of all normalized Katz centrality values equals 1. In the same way, we can normalize the centrality values for each measure of centrality. After normalizing the centrality values, we sort them in descending order, as the highest centrality value will ranked at the top of the resulting list for each measure of centrality.

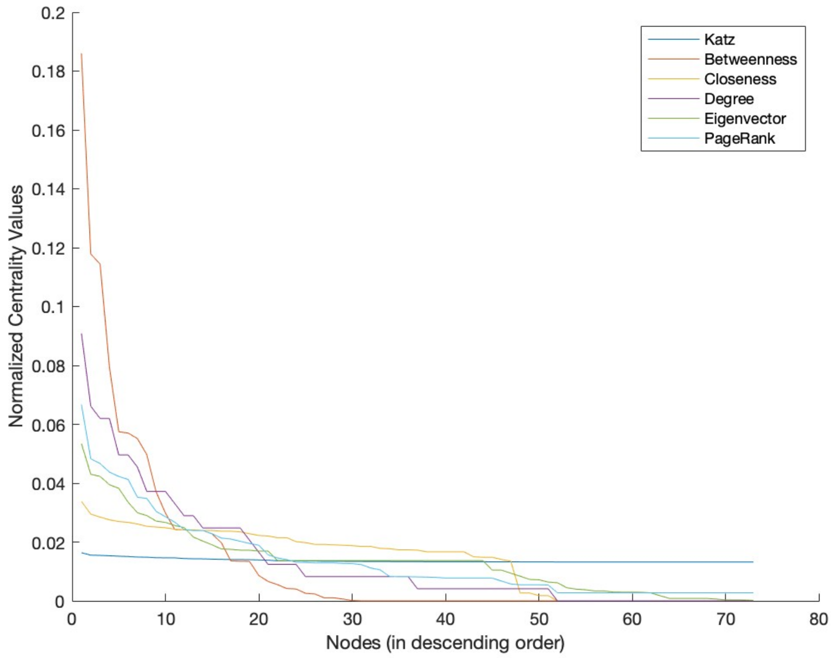

Figure 6: Normalized Centrality Measures

By comparing the centrality measures in the Figure 6, we can conclude that Betweenness and Degree centrality highlight stocks that play significant roles as intermediaries or have a high number of connections within the network. These stocks have the potential to control the flow of information and transactions, making them important players in the network dynamics.

On the other hand, Katz centrality values are relatively flat and different from the other centrality measures. This suggests that Katz centrality may not capture the influence and importance of stocks in the network as effectively as other measures. However, surprisingly, despite its low values, Katz centrality proves to be the most effective measure in identifying influential stocks. This finding suggests that there might be underlying characteristics of the network that make Katz centrality particularly suitable for capturing the true influence of stocks.

Through this analysis, we obtained the centrality of stocks for each measure. We found the three stocks with the highest centrality for each measure, which are the stocks (Prevas B, Karolinska Development B, TRATON, Nilörngruppen B, Starbreeze B, and HANZA). After that, the daily returns for each share were calculated, in addition to calculating the daily returns for the OMX stockholm 30 index, which is a stock market index that represents the performance of the 30 largest and most actively traded companies listed on the Stockholm Stock Exchange in Sweden, in order to find the correlation between the daily returns of these Stocks with daily returns for OMX, to find which centrality measure is most effective in predicting stock market trends.

1.8 Results and discussion:

Through this study was to find the centrality of a group of stocks on the Stockholm Stock Exchange through six measures of centrality and then determine its relationship with the OMX Stockholm 30 index. The three stocks with the highest centrality were selected according to each measure, then their daily returns were found and compared with the returns of the OMX Stockholm 30 index, to see which of these stocks correlated more with the returns of the Stockholm 30 Index. Through the results, it became clear that the stock ranked second in centrality according to the Katz measure (TRATON) is the stock most associated with the Stockholm 30 index with a value 0.6345 , as this stock (TRATON) is ranked third in centrality through the Eigenvector measure, which indicates that the Katz scale is more effective in predicting market trends Stock according to this study. The stock with the highest rank according to all measures of centrality (Prevas B) was not the stock most correlated with the Stockholm 30 Index with a value 0.4998 , so the stock with the second rank in centrality was looked at. Will there be a stock with a higher correlation to the Stockholm 30 Index than the stock (TRATON) if we examine all the stocks in the network, though?
In order to respond to the previous question, we took into account all of the stocks in the network. We then calculated their correlation with the Stockholm 30 index, and the findings showed that (TRATON) had the highest correlation with the OMX.

The aim of this study was to explore which of the centrality measures is most effective in predicting trends in the stock market. The results of this study show that six central measures can be used to predict stock market trends, by analyzing the centrality values and correlations, researchers can gain insights into the importance and influence of individual stocks within the network, helping them identify stocks that are more likely to follow market trends or exhibit a stronger correlation with the stock market index. For example, if the OMX index increases, we can expect that the stock with the highest correlation coefficient with OMX will also increase.

The degree centrality, which refers to the number of connections that a node (stock) has with other nodes (stocks) in the network. According to the results, the three stocks Prevas B, Karolinska Development B, and HANZA, obtained the highest values 22 , 16, and 15 respectively for the degree centrality from the rest of the stocks in the network, while the betweenness centrality of a specific node in the network is the frequency of appearance on the shortest path between two other nodes, as it indicated the results are that the stocks Prevas B, Nilörngruppen B, and Starbreeze B, had the higher values 349.5 , 221.4, and 215 respectively of betweenness centrality than the rest of the stocks in the network. While the results showed that the highest values according to eigenvector centrality, which measures the quality of a node's connection with other nodes in the network, are stocks Prevas B, Karolinska Development B, and TRATON with values 0.0534 , 0.0430, and 0.0423 respectively. While the stocks that obtained the highest values of closeness centrality, which calculates the proximity of a node to other nodes in the network, are Prevas B, Karolinska Development B, and HANZA with values 0.0054 , 0.0047, and 0.0046 respectively. The results also showed that the stocks that obtained the highest centrality through the PageRank, which measures the importance of a node based on the number of connections in other nodes in the network are Prevas B, Karolinska Development B, and HANZA with values 0.0667, 0.0483, and 0.0466 respectively. Finally, the highest centrality according to Katz centrality, which measures the importance of a node in the network based on the importance of the direct neighboring nodes, are Prevas B, TRATON, and Karolinska Development B, with values 91, 75.4, and 73.1 respectively.

We did not expect all methods to yield the same three top-ranked stocks, and we aim to avoid them being the same. The variations in the top-ranked stocks across different centrality measures can be attributed to the fact that each measure captures a different aspect of the network structure and relationship between stocks. All centrality measures calculate centrality based on different criteria and algorithms. Degree centrality counts the number of connections a stock has, while betweenness centrality considers the frequency of a stock appearing on the shortest paths between other stocks. Eigenvector centrality takes into account both the number and importance of a stock's connections, while closeness centrality focuses on the proximity of a stock to other stocks in the network. PageRank centrality considers the importance of a stock based on the number and quality of its connections. These distinct measures provide unique perspectives on centrality, resulting in variations in the rankings of stocks. These variations highlight the diverse roles and positions that stocks hold within the network, emphasizing the complexity and multidimensional nature of centrality in a stock market network.

Table 1: Centrality Measures for Stocks

	Value	Stock Name
Degree Centrality Top	22	Prevas B
Degree Centrality Second Top	16	Karolinska Development B
Degree Centrality Third Top	15	HANZA
Betweenness Centrality Top	349.5	Prevas B
Betweenness Centrality Second Top	221.4	Nilörngruppen B
Betweenness Centrality Third Top	215	Starbreeze B
Eigenvector Centrality Top	0.0534	Prevas B
Eigenvector Centrality Second Top	0.0430	Karolinska Development B
Eigenvector Centrality Third Top	0.0423	TRATON
Closeness Centrality Top	0.0054	Prevas B
Closeness Centrality Second Top	0.0047	Karolinska Development B
Closeness Centrality Third Top	0.0046	HANZA
PageRank Centrality Top	0.0667	Prevas B
PageRank Centrality Second Top	0.0483	Karolinska Development B
PageRank Centrality Third Top	0.0466	HANZA
Katz Centrality Top	91	Prevas B
Katz Centrality Second Top	75.4	TRATON
Katz Centrality Third Top	73.1	Karolinska Development B

To find out the centrality measure which is the most effective in predicting stock market trends in the network, this study relied on finding the correlation between the daily returns of the highest central shares with the daily returns of the OMX stockholm 30 index, the following table shows the results of the correlation:

Table 2: Centrality Measures and Correlation Coefficients

Centrality Measure	Top Stocks	Correlation with return of OMX Stockholm 30 Index
Degree Centrality	Prevas B, Karolinska Development B, HANZA	0.4998, 0.3616, 0.5113
Betweenness Centrality	Prevas B, Nilörngruppen B, Starbreeze B	0.4998, 0.3999, 0.2549
Eigenvector Centrality	Prevas B, Karolinska Development B, TRATON	0.4998, 0.3616, 0.6345
Closeness Centrality	Prevas B, Karolinska Development B, HANZA	0.4998, 0.3616, 0.5113
PageRank Centrality	Prevas B, Karolinska Development B, HANZA	0.4998, 0.3616, 0.5113
Katz Centrality	Prevas B,TRATON, Karolinska Development B	0.4998, 0.6345, 0.3616

The results in the table show that the use of centrality measures, especially Katz, can provide good information for predicting stock market trends. Whereas, the results of the centrality measures with all their scales indicate positive results, but the Katz scale was the most effective scale with a correlation value of 0.6345.

1.9 Bibliography

[1] Kenneth R Ahern. Network centrality and the cross section of stock returns. *Available at SSRN 2197370*, 2013.

[2] Collins Anguzu, Christopher Engström, John Magero Mango, Henry Kasumba, Sergei Silvestrov, and Benard Abola. Eigenvector centrality and uniform dominant eigenvalue of graph components. *arXiv preprint arXiv:2107.09137*, 2021.

[3] Stephen P Borgatti, Kathleen M Carley, and David Krackhardt. On the robustness of centrality measures under conditions of imperfect data. *Social networks*, 28(2):124–136, 2006.

[4] Stephen P Borgatti and Martin G Everett. A graph-theoretic perspective on centrality. *Social networks*, 28(4):466–484, 2006.

[5] Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2):163–177, 2001.

[6] Kam-Fung Cheung, Michael GH Bell, Jing-Jing Pan, and Supun Perera. An eigenvector centrality analysis of world container shipping network connectivity. *Transportation Research Part E: Logistics and Transportation Review*, 140:101991, 2020.

[7] Ettore Croci and Rosanna Grassi. The economic effect of interlocking directorates in italy: New evidence using centrality measures. *Computational and Mathematical Organization Theory*, 20:89–112, 2014.

[8] Christopher Engström. *PageRank in evolving networks and applications of graphs in natural language processing and biology*. PhD thesis, Mälardalen University, 2016.

[9] Gan Siew Lee and Maman A Djauhari. An overall centrality measure: The case of us stock market. *International Journal of Electrical & Computer Sciences*, 12(6), 2012.

[10] Ji Ma. Funding nonprofits in a networked society: Toward a network framework of government support. *Nonprofit Management and Leadership*, 31(2):233–257, 2020.

[11] Neven Matas. Comparing network centrality measures as tools for identifying key concepts in complex networks: A case of wikipedia. *Journal of Digital Information Management*, 15(4), 2017.

[12] Mark EJ Newman. The mathematics of networks. *The new palgrave encyclopedia of economics*, 2(2008):1–12, 2008.

[13] Parisutham Nirmala and Rethnasamy Nadarajan. Cumulative centrality index: Centrality measures based ranking technique for molecular chemical structural graphs. *Journal of Molecular Structure*, 1247:131354, 2022.

[14] Nirmala Parisutham and Nadarajan Rethnasamy. Eigenvector centrality based algorithm for finding a maximal common connected vertex induced molecular substructure of two chemical graphs. *Journal of Molecular Structure*, 1244:130980, 2021.

[15] Marie-Therese Puth, Markus Neuhäuser, and Graeme D Ruxton. Effective use of pearson's product–moment correlation coefficient. *Animal behaviour*, 93:183–189, 2014.

[16] Yannick Rochat. Closeness centrality extended to unconnected graphs: The harmonic centrality index. Technical report, 2009.

[17] Britta Ruhnau. Eigenvector-centrality—a node-centrality? *Social networks*, 22(4):357–365, 2000.

[18] Shamshuritawati Sharif and Maman Abdurachman Djauhari. A proposed centrality measure:: The case of stocks traded at bursa malaysia. *Modern Applied Science*, 6(10):62, 2012.

[19] Deepak Sharma and Avadhesha Surolia. Degree centrality. *Encyclopedia of systems biology*, page 558, 2013.

[20] Gilbert Strang, Gilbert Strang, Gilbert Strang, and Gilbert Strang. *Introduction to linear algebra*, volume 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[21] Inc. The MathWorks. graph.centrality. MathWorks Documentation. Accessed: 22 May 2023.

[22] Thomas W Valente, Kathryn Coronges, Cynthia Lakon, and Elizabeth Costenbader. How correlated are network centrality measures? *Connections (Toronto, Ont.)*, 28(1):16, 2008.

[23] Gaoxia Wang, Yi Shen, and Enjie Luan. A measure of centrality based on modularity matrix. *Progress in Natural Science*, 18(8):1043–1047, 2008.

[24] Justin Zhan, Sweta Gurung, and Sai Phani Krishna Parsa. Identification of top-k nodes in large networks using katz centrality. *Journal of Big Data*, 4(1):1–19, 2017.

[25] Junlong Zhang and Yu Luo. Degree centrality, betweenness centrality, and closeness centrality in social network. In *2017 2nd international conference on modelling, simulation and applied mathematics (MSAM2017)*, pages 300–303. Atlantis press, 2017.

[26] Qi Zhang, Rongxia Tang, Zhengan Yao, and Zanbo Zhang. A parallel pagerank algorithm for undirected graph. *arXiv preprint arXiv:2112.01743*, 2021.

1.10 Appendix

Table 3: The Stocks

	Name	Symbol	Currency	Sector
1	Concejo B	CNCJO B	SEK	Industrials
2	C-RAD B	CRAD B	SEK	Health Care
3	CoinShares International	CS	SEK	Financials
4	Dedicare B	DEDI	SEK	Health Care
5	Doro	DORO	SEK	Telecommunications
6	Egetis Therapeutics	EGTX	SEK	Health Care
7	Elon	ELON	SEK	Consumer Services
8	Empir Group B	EMPIR B	SEK	Technology
9	Eniro Group	ENRO	SEK	Consumer Services
10	Espisurf B	EPIS B	SEK	Health Care
11	Ferronordic	FNM	SEK	Industrials
12	Formpipe Software	FPIP	SEK	Technology
13	HANZA	HANZA	SEK	Industrials
14	Havsfrun Investment B	HAV B	SEK	Financials
15	Infant Bacterial TherapeuticsB	IBT B	SEK	Health Care
16	Immunovia	IMMNOV	SEK	Health Care
17	Mendus	IMMU	SEK	Health Care
18	Infrea	INFREA	SEK	Industrials
19	IRRAS	IRRAS	SEK	Health Care
20	Image Systems	IS	SEK	Industrials
21	Isofol Medical	ISOFOL	SEK	Health Care
22	Karolinska Development B	KDEV	SEK	Health Care
23	Lammhults Design Group B	LAMM B	SEK	Consumer Services
24	Maha Energy	MAHA A	SEK	Energy
25	Malmbergs Elektriska B	MEAB B	SEK	Industrials
26	Midway B	MIDW B	SEK	Industrials
27	Moberg Pharma	MOB	SEK	Health Care
28	Moment Group	MOMENT	SEK	Consumer Services
29	Micro Systemation B	MSAB B	SEK	Technology
30	Midsona B	MSON B	SEK	Consumer Goods
31	Medivir B	MVIR B	SEK	Health Care
32	NAXS	NAXS	SEK	Financials
33	Nelly Group	NELLY	SEK	Consumer Services
34	NGS Group	NGS	SEK	Industrials
35	Nanologica	NICA	SEK	Health Care
36	Nilörngruppen B	NIL B	SEK	Consumer Services
37	NOVOTEK B	NTEK B	SEK	Technology
38	Oncopeptides	ONCO	SEK	Health Care
39	Oscar Properties Holding	OP	SEK	Real Estate
40	Ortivus B	ORTI B	SEK	Health Care
41	Orexo	ORX	SEK	Health Care
42	Projektengagemang Sweden B	PENG B	SEK	Industrials
43	Pierce Group	PIERCE	SEK	Consumer Services
44	PION Group B	PION B	SEK	Industrials
45	Precise Biometrics	Prec	SEK	Technology

46	Prevas B	PREV B	SEK	Technology
47	Profilgruppen	PROF B	SEK	Basic Materials
48	Q-Linea	QLINEA	SEK	Health Care
49	Qliro	QLIRO	SEK	Financials
50	Railcare Group	RAIL	SEK	Industrials
51	Readly International	READ	SEK	Technology
52	Rizzo Group B	RIZZO B	SEK	Consumer Services
53	Saniona	SANION	SEK	Health Care
54	Sensys Gatso Group	SENS	SEK	Technology
55	Senzime	SEZI	SEK	Technology
56	Solid Försäkringsaktiebolag	SFAB	SEK	Financials
57	SinterCast	SINT	SEK	Industrials
58	Sleep Cycle	SLEEP	SEK	Technology
59	Softronic B	SOF B	SEK	Technology
60	Serneke Group B	SRNKE B	SEK	Industrials
61	Starbreeze A	STAR A	SEK	Consumer Services
62	Starbreeze B	STAR B	SEK	Consumer Services
63	Stockwik Förvaltning	STWK	SEK	Financials
64	Svedbergs B	SVED B	SEK	Industrials
65	Studsvik	SVIK	SEK	Industrials
66	TradeDoubler	TRAD	SEK	Technolgy
67	Transtema Group	TRANS	SEK	Industrials
68	Vicore Pharma Holding	VICO	SEK	Health Care
69	Vivesto	VIVE	SEK	Health Care
70	Wise Group	WISE	SEK	Industrials
71	Xspray Pharma	XSPRAY	SEK	Health Care
72	AAK	AAK	SEK	Consumer Goods
73	TRATON	8TRA	SEK	Industrials

Theorem 1. *Given a matrix A with real positive eigenvalue lambda and corresponding eigenvector v then v will not change in direction when multiplied with A but only change in scale.*

Proof. The equation $A\mathbf{v} = \lambda\mathbf{v}$ holds from the definition of eigenvalues and eigenvectors. Considering the right-hand side of the equation, λ is a scalar value that multiplies the vector $\mathbf{v}$. Consequently, each component of $\mathbf{v}$, denoted as v_i, will be scaled by λ.

Expressing the vector v as $v = [v_1, v_2, ..., v_n]$, where v_i represents the ith component of v, we can observe that multiplying v by λ yields the vector $\lambda v = [\lambda v_1, \lambda v_2, ..., \lambda v_n]$. Thus, each component of v is multiplied by the scalar value λ.

The resulting vector, λv, is equal to Av, and its behavior depends on the constant λ. If λ is positive, the resulting vector λv will merely scale the vector v, causing it to become longer or shorter, while preserving its original direction.

Consequently, when the eigenvalue λ is positive, the eigenvector v only undergoes a change in scale and not in direction when multiplied by the matrix A.
Hence, the statement has been formally proven.

□