Iranian Journal of Management Studies (IJMS) 2022, 15(3): 569-593
RESEARCH PAPER

A Second-Order Hierarchical Clustering of Cryptocurrencies

Hojjatollah Sadeqi*

Department of Accounting and Finance, Faculty of Humanities and Social Sciences, Yazd University, Yazd, Iran

(Received: March 6, 2021; Revised: September 27, 2021; Accepted: October 3, 2021)

Abstract
The clustering of cryptocurrencies as an emerging field in investment management is the main topic of this research. Applying the information-based distance matrices, we clustered the 30 most valuable cryptocurrencies. Then, we identified the most influential clustering by the concept of Minimum Spanning Tree (MST) and the centrality measures of graph theory. A second-order clustering, which is defined as the clustering of hierarchical clusterings, was applied to cluster 56 dendrograms. Using the most influential clustering, we identified the main clusters of cryptocurrencies and sub-clusters. The results showed that the clustering composition of cryptocurrencies changed at the period I (before COVID-19) and II (pandemic time).

Keywords: hierarchical clustering, minimum spanning trees, entropy, cryptocurrencies.

1. Introduction

I submit that the cybernetics of observed systems we may consider to be first-order cybernetics; while second-order cybernetics is the cybernetics of observing systems (Von Foerster, 2002).

Inspired by the cybernetics of cybernetics of Von Foerster (2002), we introduce the dendrogram of dendrograms and organize its building blocks in the framework of this research. Therefore, this study discusses at least four different issues: Essentially, cryptocurrencies as an emerging subject in the financial world form the central topic of this paper, while from the viewpoint of methodology, hierarchical clustering is the primary method of this research. In hierarchical clustering, the techniques of computing the distance between entities are of particular importance. Therefore, the information-based method of distance measurement is another question that this research addresses. The fourth issue is the application of Minimum Spanning Tree to identify the most influential clustering. The methodological innovation of this paper is proposing the concept of second-order hierarchical clustering, which we have defined as the clustering of clusterings.

1.1. Cryptocurrencies

Cryptocurrencies, one of the applications of blockchain technology, have shifted the paradigm of finance, business, and social contracts. They have actualized Blockchain technology potentials so dramatically that Tapscott and Tapscott (2016) have called it the Blockchain revolution. Some financial studies have focused on the technical features of this phenomenon. However, some researches discussed the price and return of digital currencies and their

* **Corresponding Author, Email:** sadeqi@yazd.ac.ir

modeling (Bouri et al., 2019). The trading volume of these financial assets is increasing steadily, and there is a tendency among investors and financial analysts to consider the risk-return trade-off of cryptocurrencies as a proper instrument of finance and investment. Some studies have applied machine learning techniques to predict the price or return of these assets (Mallqui & Fernandes, 2019). Other financial studies have examined the interaction of cryptocurrencies with other economic variables, including traditional currencies (Andrada-Félix et al., 2020), precious metal commodities (Rehman & Vinh Vo, 2020), crude oil prices (Okorie & Lin, 2020), and equity funds (Kristjanpoller et al., 2020). Following this body of knowledge, this study focuses on the hierarchical clustering of selected cryptocurrencies to improve financial decisions such as portfolio selection, risk hedging, and pairs trading.

Creating a portfolio that includes cryptocurrencies is important to some investors because of their risk characteristics and returns. Hence, the formation of such portfolios has been studied by different methods. In one study, the classical method of portfolio formation was applied based on Markowitz's approach to 500 cryptocurrencies (Brauneis & Mestel, 2019). The Black and Litterman model (Black & Litterman, 1992) has also been used as a method of portfolio diversification in the study of cryptocurrency risk (Platanakis & Urquhart, 2019). In another study, the top ten cryptocurrencies were used to form the basis of the portfolio (Liu et al., 2019). One of the methods that shows the structural relationships of cryptocurrencies with each other and with other economic variables is the Copula functions (Boako et al., 2019; Tiwari et al., 2019). However, it appears that probably the most important challenge in studying cryptocurrencies is how to calculate and estimate their risk. Various methods of risk estimation including conditional risk criteria and tail study of statistical distributions have been used in studying cryptocurrencies (Bouri et al., 2020; Liu et al., 2020; Xu et al., 2021; Zhang et al., 2021).

1.2. Hierarchical Clustering

As a classical method of machine learning, hierarchical clustering has a wide variety of applications. Financial decision-makers also utilize this method in their analysis (Khedmati & Azin, 2020). However, this method, as a method of knowledge discovery, has its own strengths and weaknesses (Cai et al., 2014). The end product of hierarchical clustering is the various dendrograms that show the metric position of the clustered entities. Two essential issues mutually generate various clusters (and dendrograms): the clustering method and the method of measuring the distance between entities. Therefore, if there are n clustering methods and m distance measurement methods, the desired entities can be clustered into $n * m$ different clusterings. The critical question is, "Which of all these hierarchical clusterings should we take as the basis for our decision?" In the clustering literature, there are methods such as tanglegram to calculate the degree of convergence and divergence of hierarchical clusters (Scornavacca et al., 2011). However, the superiority of a dendrogram to others is still a controversial argument. Applying the Minimum Spanning Tree of graph theory, this research could support the decision-makers to identify the most suitable clustering.

1.2.1. The Problem of Hierarchical Clustering

More specifically, in any clustering, there are several methods for calculating the distance between entities and creating clusters. Therefore, with the same data, multiple clusters are obtained. By limiting the subject to entities that are of the time series type and by limiting the method to hierarchical clustering, this problem is formulated as follows: which method of calculating time series from each other and which method of creating hierarchical clusters does it take precedence over others? While methods have been proposed for comparing clusters, the novelty of this paper is that it compares clusters with each other and creates a hierarchy of clusters. In this hierarchy, the position of the clusters is measured from each other and we can find out which clusters are similar. But there are also several methods for clustering hierarchies whose entities are clustered. This leads into a vicious circle. To get out of this vicious circle, we have used the concepts of graph theory. If we consider each of the resulting clusters as vertices of a graph, then we can analyze the resulting graph in different ways. One of these methods is to calculate the criteria for the centrality of graph vertices. A range of centrality measures allows us to identify the most significant clustering.

1.2.2. Comparison of Clusterings

Different algorithms and methods create clusters that are not necessarily the same. By comparing 22 similarity indices, a study set out to compare clusterings (Albatineh et al., 2006). These indicators compared different clustering methods and showed the degree of similarity or dissimilarity with a quantity between 0 and 1. Another study examined the relationship between some of these indicators (Warrens, 2019). However, Van der Hoef and Warrens (2019) focused on similarity indicators based on information theory. On the other hand, some software packages have made it possible to measure these indicators in applied studies. The *CluSim* package (Gates et al., 2019) in Python and the *dendextend* package (Galili, 2015) in R allow the calculation of several similarity criteria. In this study, we used the *dendextend* package. However, these indicators compare clusters in pairs and measure their similarity. In this research, we go one step further and use a graphic design which is headed by clusters. In this graph, in addition to representing the interrelationships of clusters, the importance of each clustering is calculated based on the criteria of centrality.

1.2.3. Clustering of Time Series

In some fields of study, such as finance and economics, time series are the basis of much research. In addition to classic econometric methods, various machine learning methods, especially time series clustering, can provide useful information to decision makers. In one study (Warren Liao, 2005), the time series clustering literature was examined in detail. The author, while reviewing previous research in terms of algorithms and performance criteria for clustering and the basis for defining the similarity of time series, divided them into three groups, depending on whether they work directly with raw data in the time or frequency domain, indirectly with features extracted from raw data, or indirectly with the models on raw data. In another study (Aghabozorgi et al., 2015), research over a decade in various fields, from finance to biology, based on time series clustering, was reviewed. The authors of this paper referred to different types of classical clustering methods and newer methods such as fuzzy clustering methods. The study revealed that one of the most important decisions in time series clustering is to choose the clustering method.

1.2.4. Hierarchical Structure of Cryptocurrencies

The study of hierarchical structures in financial markets began with the important article of Mantegna (1999). In this article, the author discusses the concept of hierarchical structure in stock markets, according to which, the relationships between the shares of a market can be understood in the form of an overview. According to the Web of Science website, this article has been cited 1065 times so far, of which 873 have been research articles.

The analysis of bibliographic information of these 873 research articles using the *bibliometrix* R package (Aria & Cuccurullo, 2017) shows that the most cited articles

these 873 have dealt with the application of hierarchical clustering in financial markets (Bonanno et al., 2004; Mantegna, 1999).

On the other hand, we see the most important concepts that have been studied in articles in Figure 1.

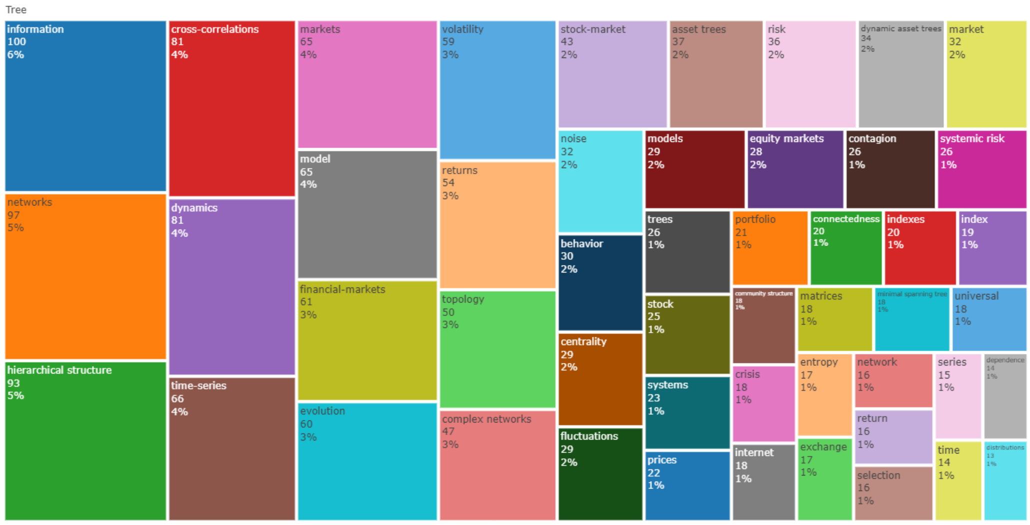

Figure 1. The Main Concepts of Hierarchical Structure of Financial Markets

Some recent studies have focused on the hierarchical structure of the cryptocurrency market. The hierarchical structures of the cryptocurrency market show that Bitcoin and Etherium have a well-established leadership in this market. On the other hand, using the cross correlation of prices and the related Minimum Spanning Tree, a number of homogeneous clusters have been reached (Song et al., 2019). In another study, using the same hierarchical method, the centrality of Bitcoin in the cryptocurrency market was investigated. The formation of the Minimum Spanning Tree in this study is the shock transmission of the Bitcoin price and its effect on other cryptocurrencies that were measured with vector autoregressions (Zięba et al., 2019). In another study, cryptocurrencies were analyzed using a hierarchical structure based on their price correlation, and the results were reviewed using a Random Matrix (eigenvalues analysis) method (Stosic et al., 2018). The structure of communities and the dynamics of price correlation in the cryptocurrency market were studied in another study. In this study, the collective behavior of 119 cryptocurrencies in 2017 and 2018 was analyzed (Chaudhari & Crane, 2020).

In the most of these studies, the distance between cryptocurrencies has been calculated based on price correlation. Due to the fact that information-based distances may lead to different results, this study, in addition to generalizing the previous studies (to the two periods before and after the COVID-19), methodologically relies on information-based distances.

1.3. Information-Based Distance

The calculation of the distance matrix is an essential step in hierarchical clustering. Distance computation depends on the nature of the objects and their characteristics (Deza & Deza, 2013). In the hierarchical clustering of cryptocurrencies, the objects are time series of prices and returns. Despite all the challenges, the correlation coefficient is still the dominant method for calculating the distance of time series objects. Some studies have proposed alternative algorithms like dynamic time warping (DTW) that do not rely on the correlation coefficient (Giorgino, 2009).

One way to define the distance of time series objects is to calculate the distance matrix based on the concepts of information theory (Kraskov & Grassberger, 2009). The mutual information of the two variables originates from the concept of entropy and shows a kind of similarity between the two variables. An appropriate transformation can turn this similarity criterion into a distance measure (Hu et al., 2017). In most financial clusterings, the basis for measuring distance has been the correlation coefficient. However, this study uses methods based on information theory as a complementary method and compares the results. Nonetheless, the calculation of distance matrices for the time series of cryptocurrencies by all distance methods and their comprehensive comparison with each other requires independent research.

The use of information-based distances in financial clustering has been discussed in some studies, but in the study by Guo et al. (2018), different methods for calculating these distances are examined.

1.4. Minimum Spanning Tree

The Minimum Spanning Tree (MST) is a classical concept of graph theory to filter the large graphs and remove extra edges. Therefore, some financial studies have applied MST to recognize central stocks of large stock networks (Coletti, 2016). A stock network is a case of financial network that represents the relationship of financial objects such as financial assets, stocks, currencies, and stock indexes. Applying the algorithms of Minimum Spanning Tree and other filtering techniques of large financial networks, we can calculate the centrality measures to identify the most influential nodes in the filtered financial networks (Jang et al., 2011; Jo et al., 2018; Tabak et al., 2010). In this study, we generated a graph whose vertices were hierarchical clusterings constructed with several techniques. Then, we filtered this network of clusterings with Prim's algorithm to make their Minimum Spanning Tree. After calculation of the centrality measures of this MST, we identified the most influential clustering to make a more reliable financial decision. We investigated the first-order (or raw) hierarchical clusterings by the concepts of graph theory in addition to the hierarchical clustering theory. This process is a second-order hierarchical clustering to determine the superior clustering. MST is one way to filter complete graphs. In addition to this method, other methods such as Planar Maximally Filtered Graph have been used in some studies as a basis for filtering complex financial networks (Tumminello et al., 2005; Tumminello et al., 2007).

2. Data and Methodology

2.1. Time Series of Cryptocurrencies

Bitcoin has a high reputation as the most well-known cryptocurrency and has been the subject of numerous studies (Aggarwal, 2019). The popularity of bitcoin has led to misconceptions that concepts such as blockchain and cryptocurrencies are the same as bitcoin (Merediz-Sola & Bariviera, 2019). There are hundreds of cryptocurrencies generated by blockchain technology. Therefore, to narrow the scope of research, we have selected the top 30 cryptocurrencies and their daily price data according to their market shares and trading volumes. We downloaded it from *Yahoo!Finance* and calculated their daily logarithmic returns using Equation 1.

$$r_t = \ln(p_t) - \ln(p_{t-1}) \tag{1}$$

The *quantmod* package of R, developed for quantitative financial analytics and Quants, provides us to retrieve data from *Yahoo!Finance* (Ryan & Ulrich, 2020). Some research has shown that the COVID-19 pandemic has had a significant impact on financial markets, including cryptocurrencies and investor behavior (Corbet et al., 2020; Ortmann et al., 2020; Umar & Gubareva, 2020). Some studies have examined the efficiency and stability of the cryptocurrency market before and after the COVID-19 pandemic. They show that these markets have been less stable after this event (Lahmiri & Bekiros, 2020; Mnif et al., 2020). Accordingly, given the occurrence of the COVID-19 outbreak, we divided the research into two time periods of the last six months of 2019 (period I) and the first six months of 2020 (period II), and repeated the research steps for both periods. According to the descriptive statistics of observations (daily returns of selected cryptocurrencies), the average daily returns of cryptocurrencies in the period II compared to the period I has increased.

2.2. Calculation of Distance Matrices

After calculating the returns, the distance matrix was compiled based on them. In the computation of the distance between cryptocurrencies, we used two approaches of correlation-based and information-based distances. In information-based distances, the concept of entropy plays a critical role. The entropy of the random variable X was determined as follows by Equation 2 (Hu et al., 2017):

$$\boldsymbol{H}(X) = -\sum_{x \in \Omega} p(x) \log_2 p(x) \tag{2}$$

In the Rényi method for calculating entropy, there is also the parameter α. By setting this parameter, we can manipulate the weight of extreme observations. The Extreme value theory approach to risk measurement focuses on the observations at the tail of the statistical distribution (Rasmussen, 2014). If we set the parameter α equal to one, the Rényi's entropy and Shannon's entropy are equivalent. In other words, Rényi's entropy is a generalized form of Shannon's entropy (Amigó et al., 2018). How to calculate the Rényi's entropy for variable X is as in Equation 3 (for a detailed exploration of the features and applications of Rényi's entropy, see Principe, 2010).

$$\boldsymbol{H}_{\alpha}(X) = \frac{1}{1-\alpha} \log_2 \sum_{x \in \Omega} p(x) \tag{3}$$

For the time being, in this study, we ignored the effect of α and its different values on the entropy value. After measuring the entropy of the variables, we calculated the mutual information of the two variables.

Mutual information, as a measure for estimating the statistical dependence of variables, shows the amount of common information within the two variables X and Y (Bossomaier et al., 2016). Therefore, mutual information indicates the similarity of variables and consequently, it has the same application as correlation coefficient in clustering and classification. If we convert this similarity measure into a distance metric with an appropriate transformation, the obtained distance matrix can be an alternative of the correlation-based distance matrices applied in the clustering process. In this study, we examined several methods for converting mutual information into a distance metric and demonstrated the application of such information-based distance criteria in cryptocurrencies analysis.

Mutual information represents a relationship between the variables that conduces to the correlation coefficient in certain circumstances (Song et al., 2012). If knowing one variable reduces the entropy of the other variable, there is common information between them. The independence of those two variables occurs when knowing one of them tells us no truth about the other. That means the mutual information of the two variables in the case of independence is zero. On the other hand, this definition shows that mutual information describes a symmetrical and non-directional relationship (Batina et al., 2011).

$$\boldsymbol{I}(X:Y)=\boldsymbol{H}(X)-\boldsymbol{H}(X|Y)=\boldsymbol{H}(Y)-\boldsymbol{H}(Y|X) \quad (4)$$

In other words, mutual information equals the sum of entropies minus the common entropy and expresses a particular instance of Kullback–Leibler divergence. It is always a nonnegative number (Cover & Thomas, 2006).

$$\boldsymbol{I}(X:Y)=\boldsymbol{H}(X)+\boldsymbol{H}(Y)-\boldsymbol{H}(X,Y) \quad (5)$$

Equation 6 shows the probabilistic form proposed to calculate the mutual information (Steuer et al., 2005).

$$\boldsymbol{I}(X:Y)=\sum_{x\in\Omega_x,y\in\Omega_y} p(x,y)\log_2\frac{p(x,y)}{p(x)p(y)} \quad (6)$$

After the definition of mutual information as a criterion of similarity between the two variables, it is essential to define distance measurement criteria. The easiest way to convert mutual information into a distance metric is to use the complementary method. Complementary information is called “Variation of Information” (or Sum Distance), which is defined as in Equation 7 (Meila, 2003).

$$d_{sum}(X,Y)=\boldsymbol{H}(X|Y)+\boldsymbol{H}(Y|X) \quad (7)$$

The normalized form of the sum distance is called “Shared Information Distance” (Li, 2006).

$$d_{shared}(X,Y)=\frac{\boldsymbol{H}(X|Y)+\boldsymbol{H}(Y|X)}{\boldsymbol{H}(X,Y)} \quad (8)$$

Another method is to use the maximum operator. In this method, unlike the shared information distance, we decompose the variation of information into its components and then consider their maximum as a distance metric (Ming et al., 2004).

$$d_{max}(X,Y)=\max\left(\boldsymbol{H}(X|Y),\boldsymbol{H}(Y|X)\right) \quad (9)$$

The normalized form of max distance is defined as in Equation 10 (Ming et al., 2004).

$$d_{nmax}(X,Y)=\frac{\max\left(\boldsymbol{H}(X|Y),\boldsymbol{H}(Y|X)\right)}{\max\left(\boldsymbol{H}(X),\boldsymbol{H}(Y)\right)} \quad (10)$$

To calculate the entropy and mutual information of variables and distances based on information, we used the *infotheo* R package and its functions. To calculate the variable entropy, first the data discretization process is performed to determine their statistical distribution. In this study, we used an experimental method for data discretization (Meyer, 2014).

In the *infotheo* R package, discretization methods are programmed based the equal frequencies or equal width binning algorithm. In this study, the input of the function was set to the discretization method based on equal frequency and the number of bins was equal to one third of the number of observations (as the default value of the function). Checking the effect of the number of bins on estimating the distribution of observations can be considered as a separate issue. For comparison, we also calculated correlation-based distances using Pearson, Spearman, and Kendall methods of correlation. In converting correlations to distance metrics, we used the following conversion of Equation 11.

$$d_{cor}(X,Y) = \sqrt{2*(1-\rho_{X,Y})} \tag{11}$$

2.3. Different Methods of Hierarchical Clustering

After computation of the distance matrices by the mentioned methods, we had to select the clustering algorithm. There are two main approaches proposed to find the hierarchy of clusters in our observations: The bottom-up strategies and vice versa.

The bottom-up strategies (also called agglomerative algorithms) start discovering clusters from the bottom of the observations. They integrate the selected groups of observations into a new cluster recursively at the next level. This process develops a cluster at the next level with one less cluster because of integrating lower-level clusters (Dubitzky et al., 2013).

When deciding to merge two clusters from one level to the next, calculating the distance of the clusters from each other is a crucial step. There are several ways to measure the distance of clusters, the most famous of which are Single linkage, Complete linkage, Group average, Mcquitty method, Median method, Centroid method, Ward.D and Ward.D2 methods. The number of dendrograms was determined by multiplying the number of methods (7 methods) we used to calculate the distance matrix by the number of clustering methods (8 methods). Therefore, there were 56 hierarchical clusterings to analyze in the next step.

2.4. A Minimum Spanning Tree of Dendrograms

Dendrograms are the end product of hierarchical clustering, and by comparing them in pairs, we determine the degree of their convergence or divergence. In comparing dendrograms, we can use the tanglegram method and their degree of entanglement, as well as the correlation between dendrograms. These methods are coded using the *dendextend* package in R. There are several methods to calculate the correlation (or the similarity) of hierarchical clustering trees such as the Cophenetic method (Sokal & Rohlf, 1962), FM index, and Baker's Gamma correlation coefficient.

One approach to identify the most influential dendrograms is to apply graph theory. A graph is a set of vertices and edges that show the relationship between those vertices. Thus, in our graph, the vertices were 56 dendrograms (hierarchical clustering trees) and the edges represented the distance of those dendrograms, which were calculated based on the cophenetic correlation coefficient. The cophenetic correlation coefficient as a measure of similarity between two dendrograms is a number between −1 to 1. Sokal and Rohlf (1962) defined this method, and there is a function in the *dendextend* R package that calculates it, which can be used to obtain the dendrogram correlation matrix.

We again used Equation 11 to convert the correlation (similarity) matrix of the dendrograms to their distance matrix. Using this matrix as an adjacency matrix of a graph, we arrived at a complete graph of order 56, each edge of which showed the distance from one dendrogram to others, for a total of 1540 edges.

$$\frac{N(N-1)}{2} = \frac{56*55}{2} = 1540$$

Next, we filtered this relatively large graph via the MST algorithm. As a method of filtering the edges of a connected, weighted undirected graph, a Minimum Spanning Tree (MST) is a subset of the edges that connects all the nodes of that network. It has the minimum possible total edge weight without any cycles.

There is a function in the *igraph* R package (Csardi & Nepusz, 2006) that computes the minimum spanning tree of a graph based on the Prim's algorithm. Then, using the centrality measures proposed in graph theory, we identified the most influential dendrograms in this MST or network of hierarchical clustering.

There are numerous studies about centrality metrics in the graph theory literature, and every day, graph theory researchers introduce new criteria. This study focused on the three most well-known measures of centrality and identified the most influential dendrogram in order of priority. These criteria are delineated by Golbeck (2013) as follows.

(1) Degree: The degree of a vertex in a graph (a node in a network) is its most essential structural property. It is simply defined as the number of its adjacent edges. How many vertices are connected to that vertex? The edges of a graph represent these connections.

(2) Closeness: This measure indicates how far, on average, one vertex is from the other vertices of the graph. And how many steps must be taken to reach those vertices? The fewer these steps, the greater the value of this measure of centrality.

(3) Betweenness: This centrality criterion indicates how many times a vertex is located between other vertices. In finding the shortest path between two vertices, we cross the other vertices. The more a vertex is present in the shortest paths, the more remarkable that vertex is. That is, it represents a powerful intermediary or bridge between other vertices.

3. Results

Downloading the relevant data of the selected cryptocurrencies from *Yahoo! Finance*, we calculated the time series of their returns.

3.1. A Dendrogram of Dendrograms

Following the calculation of the distance matrix, we made 56 dendrograms of Cryptocurrencies. A methodological problem in applying these dendrograms for financial decisions is their divergence. In other words, each of these dendrograms gives a distinctive result, and none of them are the same.

For example, Figures 2 and 3 show the circular dendrograms obtained from two different methods: Method A and Method B. As shown in Method A, the observations are grouped into two clusters: the first cluster contains only one cryptocurrency and the second cluster includes other cryptocurrencies. Such clustering does not provide financial analysts with useful information for decision-making.

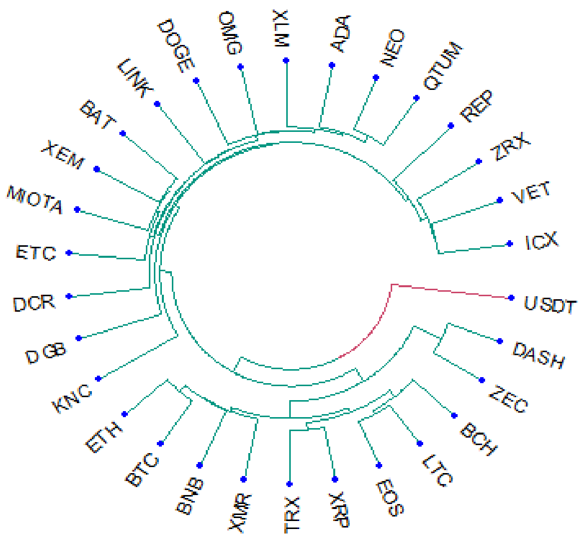

Figure 2. The First Dendrogram: Pearson-based Distance and Ward.D Method

However, in Method B, we are faced with two relatively balanced clusters. Therefore, there were no similarities between the dendrograms.

The degree of similarity or distance of these dendrograms can be determined by the methods mentioned in Section one.

Figure 4 and Figure 5 show the dendrogram of dendrograms for the period I and II. First, the correlation of the clusters was calculated based on the cophenetic correlation, and after converting it into a distance criterion, we achieved a hierarchical clustering of hierarchical clustering. This achievement was a second-order hierarchical clustering.

As shown in Figure 4 and Figure 5, to prioritize a method, we needed to do more analysis on this dendrogram. At the same time, a vicious circle appeared: because these clusterings could also be re-clustered in several ways, the process of clustering could be repeated infinitely.

3.2. A Minimum Spanning Tree of Dendrograms

To get out of this vicious circle, one can look at the issue from outside the theory of clustering. The proposed approach is to correct this dilemma using graph theory. Graph theory, as part of discrete mathematics, has the potential to play a complementary and promotional role in many clustering problems. Figure 6 shows such a complete graph of dendrograms. A complete graph with 56 vertices and its dense edges does not seem suitable for analysis and visualization.

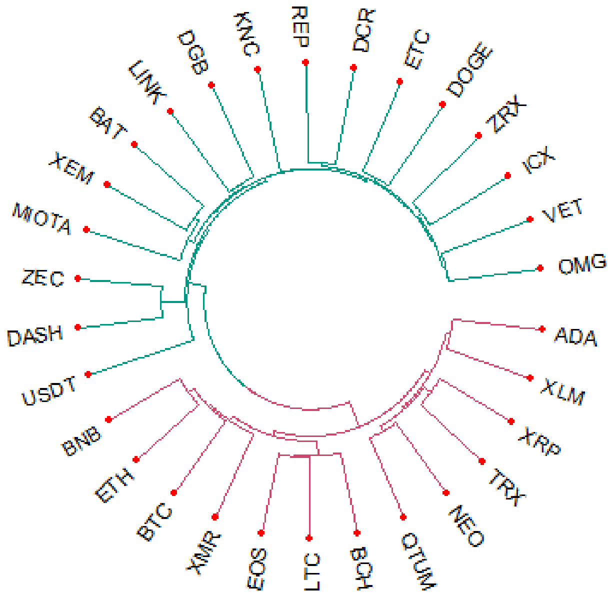

Figure 3. The Second Dendrogram: Shared Information Distance and Average Method

Considering the centrality measures, we could not distinguish dendrograms from each other. The degrees of vertices in graphs close to the complete graph were almost equal, and other criteria of centrality were not applicable to select the superior clustering.

Moreover, if the visualization of these graphs is critical to making better decisions, we need to filter them by the minimum spanning tree algorithm. Therefore, in the continuation of this study, using the distance matrix obtained for the dendrograms, we created a complete graph, and then using Prim’s algorithm, we compiled the minimum spanning tree of that graph. The minimum trees of the period I and II are shown in Figure 7 and Figure 8. As shown in Figure 7, if we consider the degree of the vertices as a decision criterion, the dendrogram proposed by Pearson-based Distance and Complete method is given priority in the period I. On the other hand, according to Figure 8, the dendrogram of Normalized Max Information Distance and Ward.D method is the most influential clustering in the period II. In the period II, the same dendrogram has the best value of the closeness centrality and the highest value of the betweenness centrality.

Figures 9 and 10 show the most influential dendrograms of the periods I and II. Various clusters of cryptocurrencies can be recognized precisely at the distances and orders that this hierarchy of clusters represents. The logic of clustering

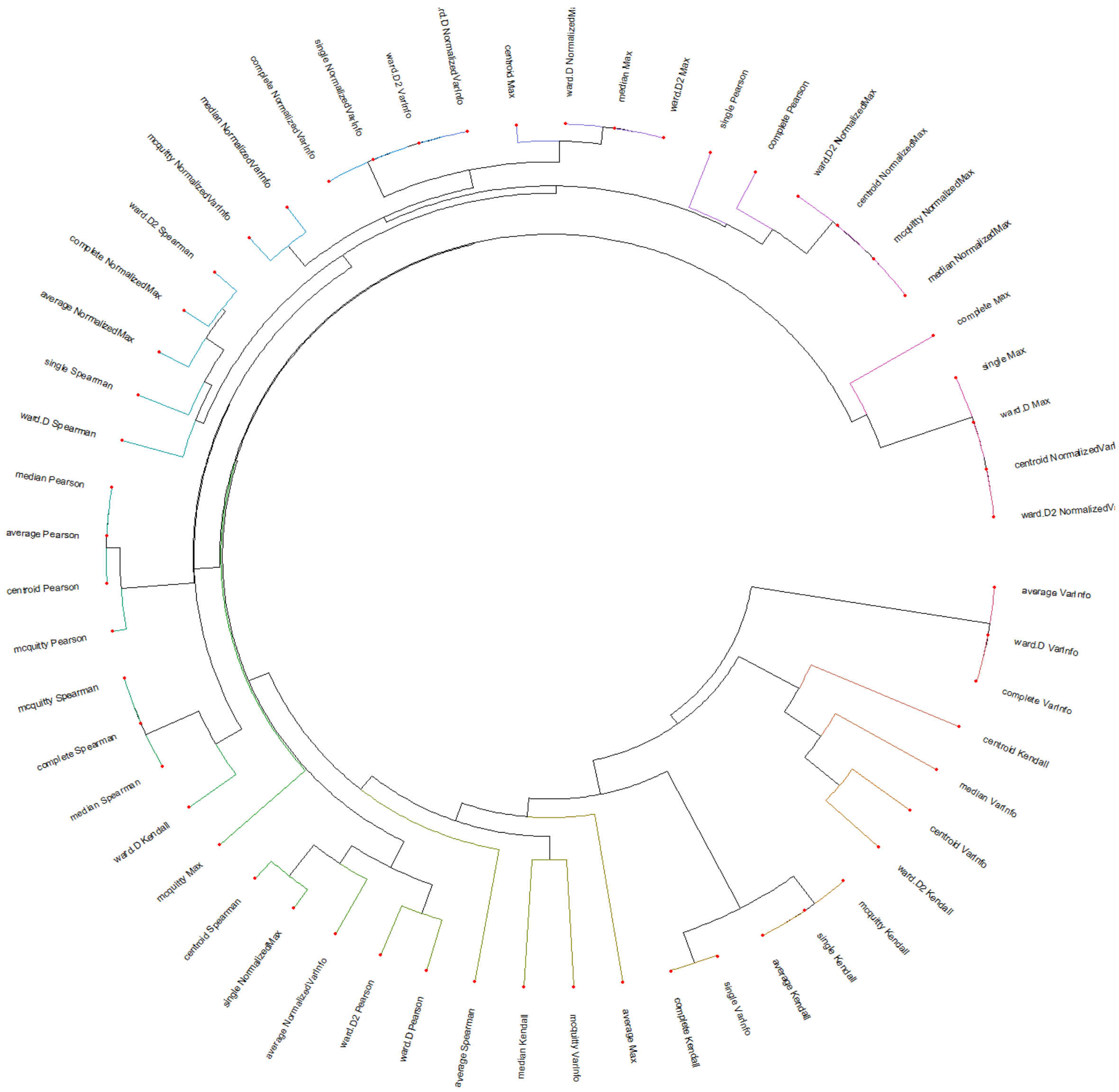

Figure 4. The Second-order Dendrogram of Hierarchical Clusterings - Period I

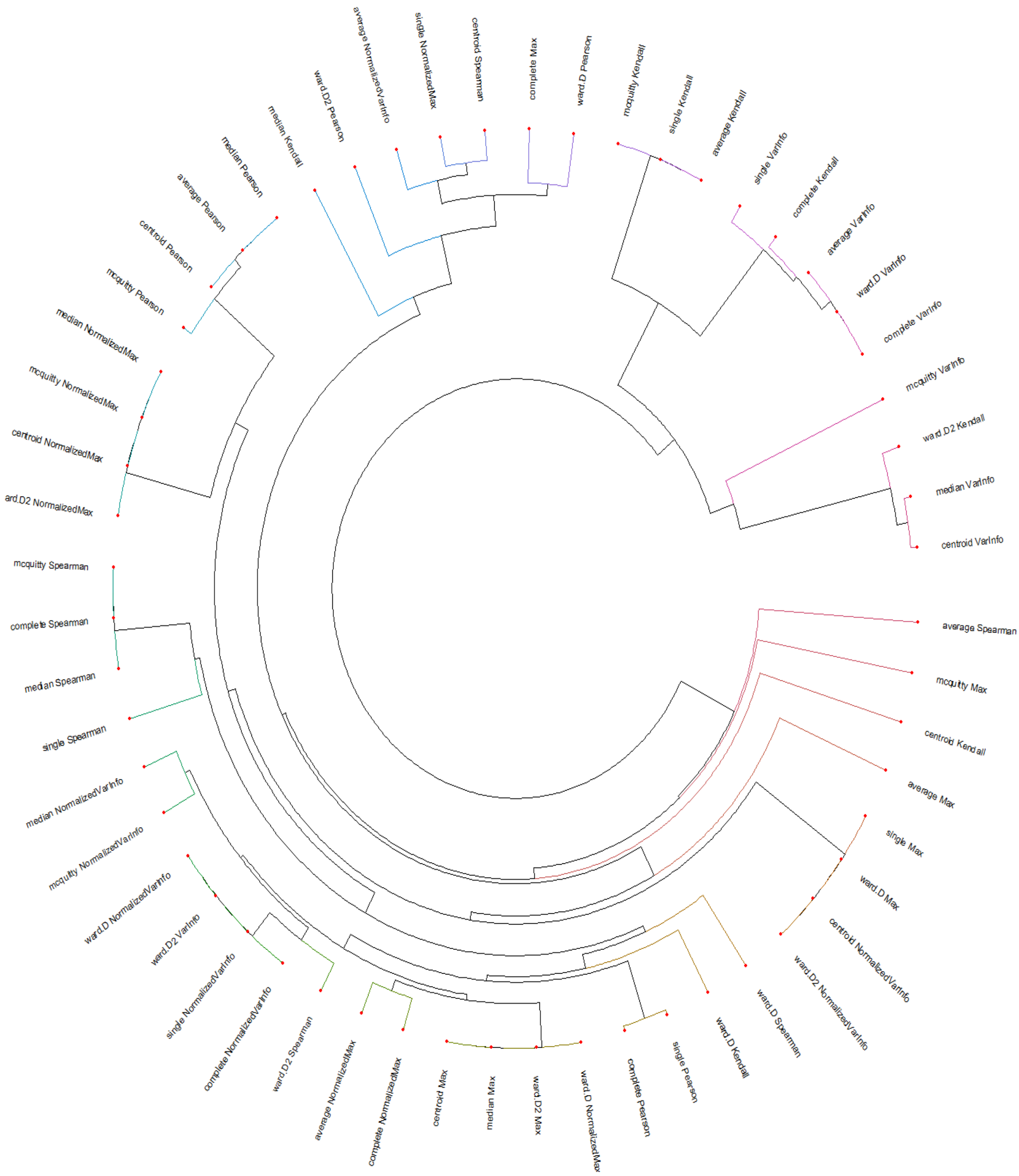

Figure 5. The Second-order Dendrogram of Hierarchical Clusterings - Period II

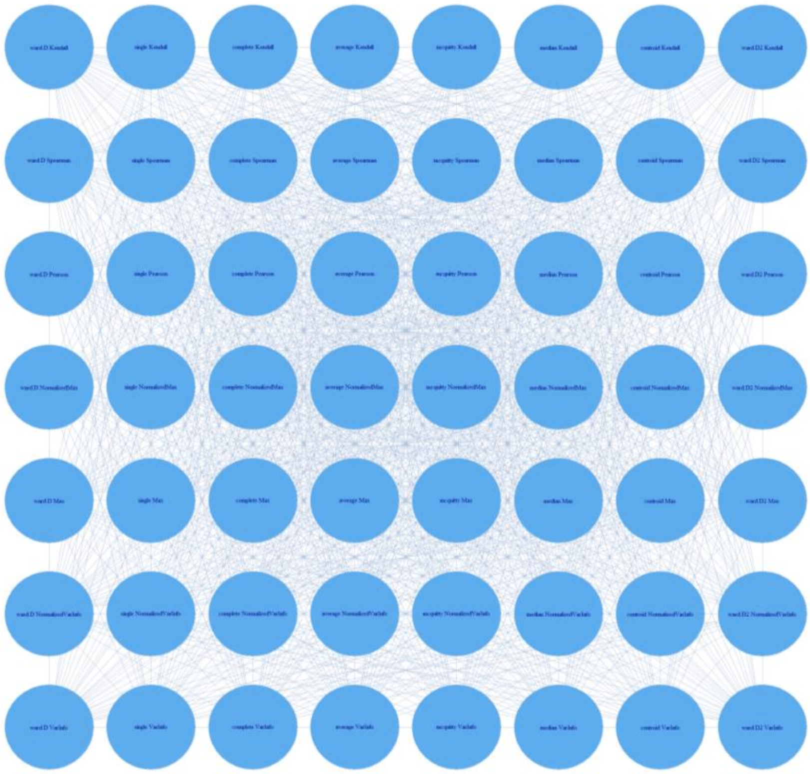

Figure 6. The Complete Graph of 56 Dendrograms

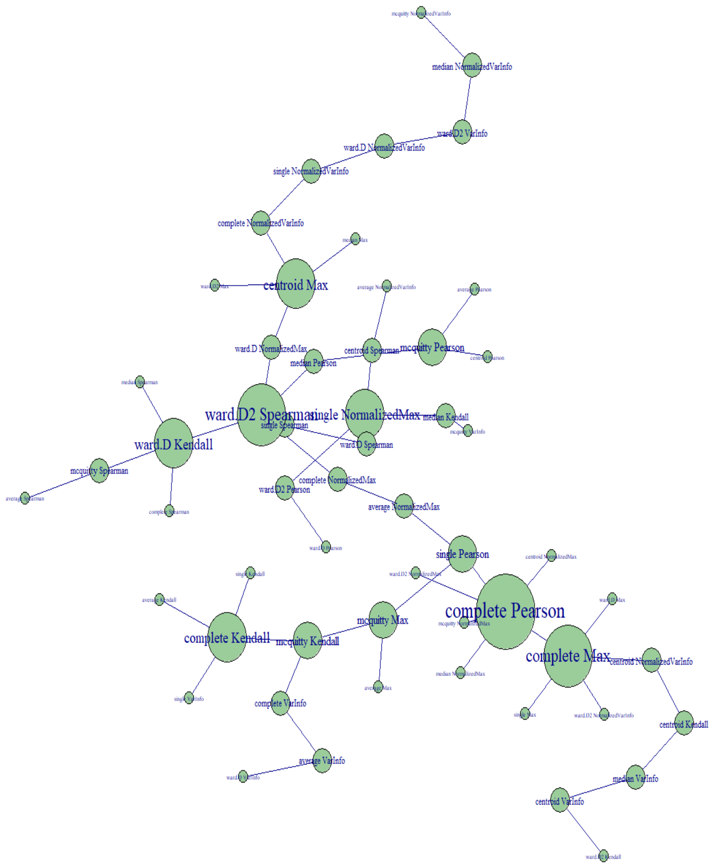

Figure 7. The Minimum Spanning Trees of 56 Dendrograms - Period I

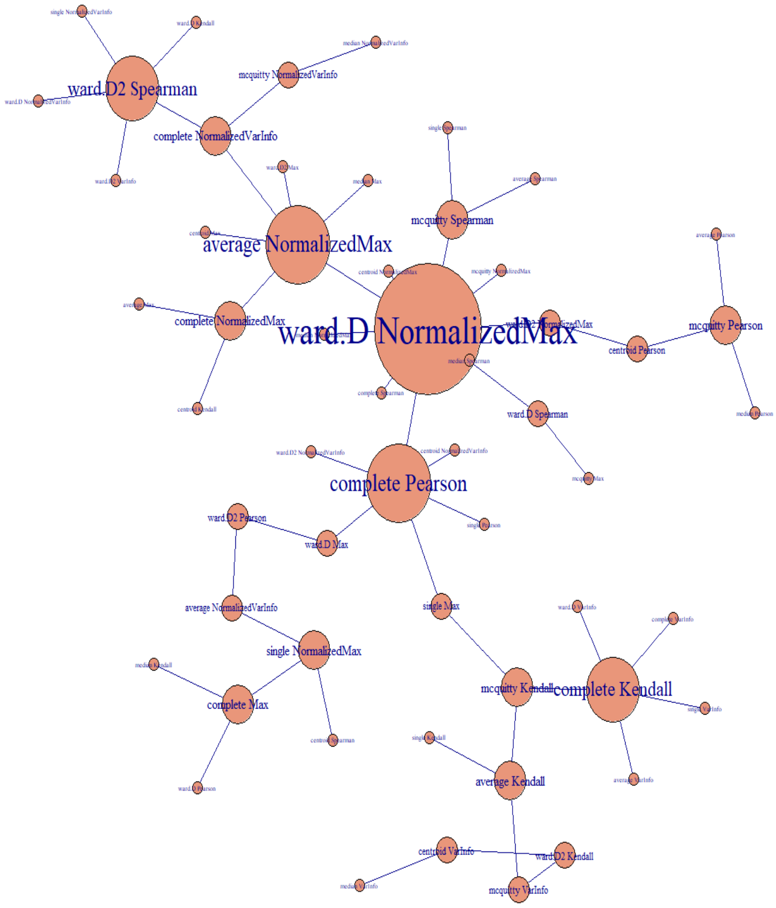

Figure 8. The Minimum Spanning Trees of 56 Dendrograms - Period II

presents an informative guideline for financial decisions: the Cryptocurrencies that are in a cluster have the shortest distance from each other. Therefore, if one of them is not possible to trade, the traders can consider the other ones clustered in the same group. In other words, the cryptocurrencies that are in a cluster can be used as substitutes for each other in portfolio selection. The principle of diversification suggests that the investors should avoid choosing a portfolio that includes cryptocurrencies in a cluster before and after COVID-19 pandemic. Based on the hierarchical clustering of Figure 10, the identified clusters in the period II have the following characteristics:

Two of the identified clusters have only one member. These cryptocurrencies have the greatest information distance from other cryptocurrencies. In other words, knowing information about KNC and USDT cryptocurrencies does not add to our knowledge of other cryptocurrencies. KNC is an Ethereum-based cryptocurrency for paying transaction costs on the Kyber Network. On the other hand, USDT (Tether) is a cryptocurrency invented to reflect the value of the U.S dollar. Tether as the digital U.S. dollar is significantly different from other cryptocurrencies. Thus, it is a unique cryptocurrency that does not fit in any cluster with any of the other cryptocurrencies.

One of the clusters has two members and includes the cryptocurrencies DASH and ZEC. DASH is an open-source cryptocurrency that is derived from Litecoin, which itself is forked from Bitcoin. DASH and ZEC are located in a cluster and form a pair. Traders can use this fact in pair trading. In such transactions, finding a pair of financial assets is the main issue. On the other hand, ZEC (Zcash) as a cryptocurrency is developed on the Bitcoin codebase. In the other cluster, we see a combination of two subclusters. One of these subclusters includes 5 cryptocurrencies: MotaCoin (MOTA), New Economy Movement or NEM (XEM), Basic Attention Token (BAT), Chainlink (LINK), and DigiByte (DGB). And another subcluster includes 8 assets: Augur (REP), Decred (DCR), Ethereum Classic (ETC), Dogecoin (DOGE), 0x (ZRX), ICON (ICX), VeChain (VET), and OMG Network (OMG). The last main cluster also consists of two subclusters. In one of these subclus-ters, 6 cryptocurrencies are located as follows: Cardano (ADA), Stellar (XLM), Ripple (XRP), TRON (TRX), Neo (NEO), and Qtum (QTUM). In the last and most important subcluster, the most prominent and oldest cryptocurrencies are placed next to each other: Bitcoin (BTC), Ethereum (ETH), Binance Coin (BNB), Litecoin (LTC), Monero (XMR), EOS (EOS), and Bitcoin Cash (BCH).

In general, the most popular currencies are placed in a cluster or a sub-cluster. Therefore, because we have clustered the objects applying information-based distances, the interpretation of this finding is so fantastic: the popular cryptocurrencies have more mutual information.

Similarly, Figure 9 shows the cryptocurrency clusters in the period I (the last six months of 2019). What is the similarity of the hierarchical clustering of cryptocurrencies before and after COVID-19? As a comparison, in the last step of this research, we compared the best dendrograms of the periods I and II. From the representation of Figure 9, it is obvious that the best dendrogram of the period I has several inconsistencies with the best clustering of the period II. The composition, shape, and position of cryptocurrencies in this dendrogram are distinct from the previous dendrogram.

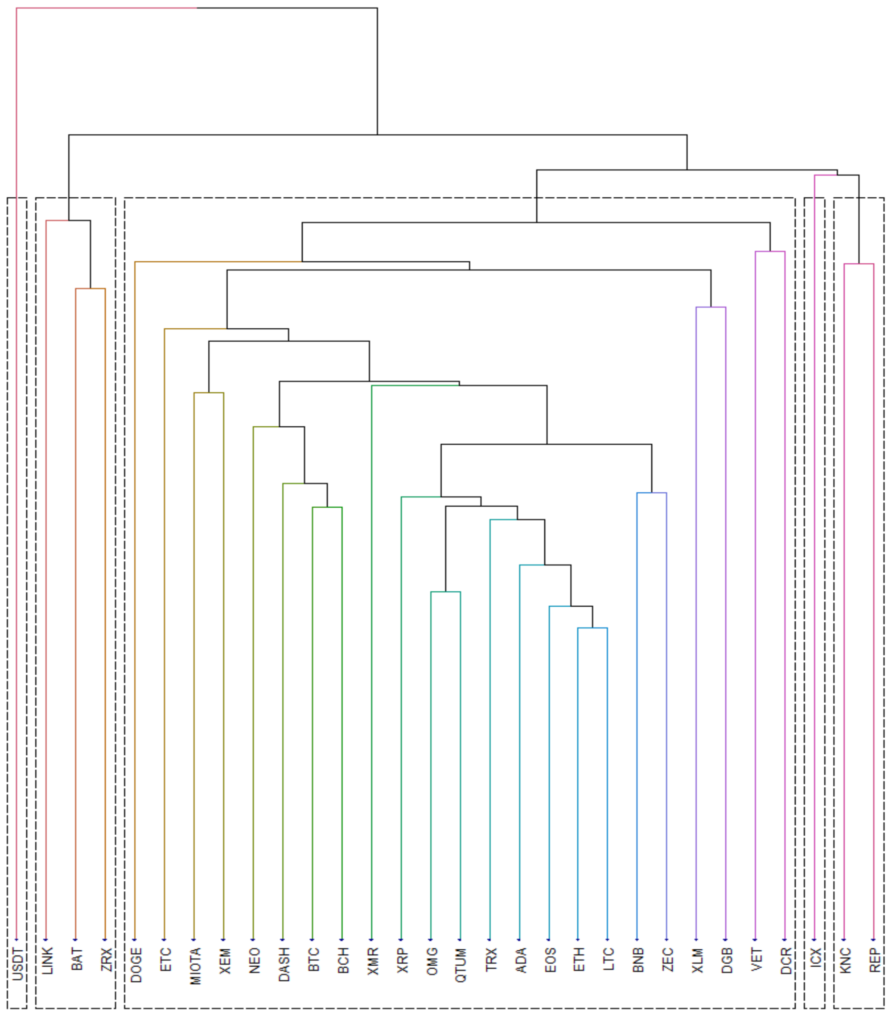

Figure 9. The Best Dendrogram - Period I: Pearson-Based Distance and Complete Method

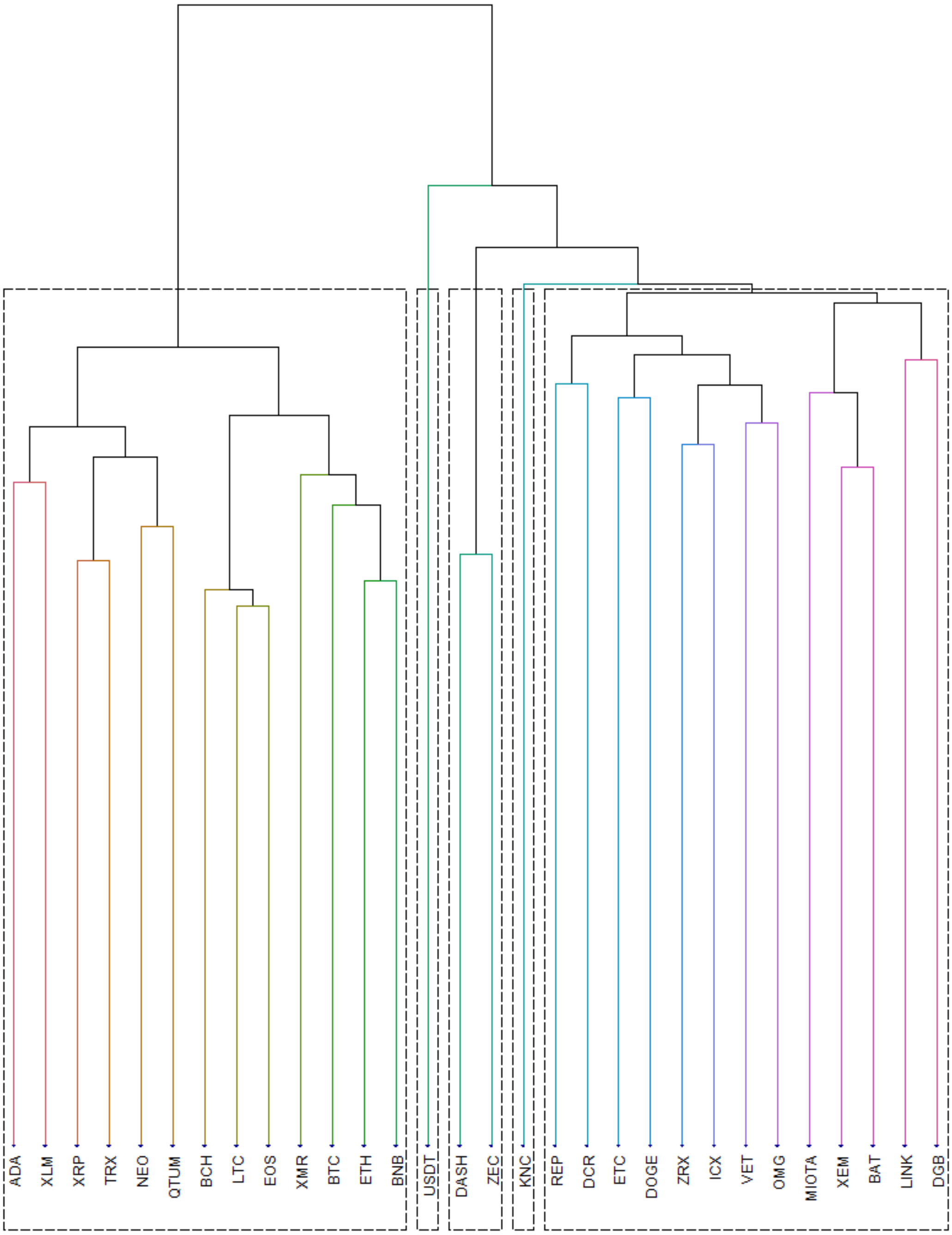

Figure 10. The Best Dendrogram - Period II: Normalized Max Distance and Ward.D Method

Tables 1 and 2 show some topological information about trees before and after COVID19.

Table 1. Centrality Measures of Dendrograms in the Complete Graph -Period I

	Degree	Closeness	Betweenness
Ward.D VarInfo	55	0.01784401	0
Single VarInfo	55	0.021449799	0
Complete VarInfo	55	0.01784401	0
Average VarInfo	55	0.01784401	0
Mcquitty VarInfo	55	0.023096885	0
Median VarInfo	55	0.01743054	0
Centroid VarInfo	55	0.018245766	0
Ward.D2 VarInfo	55	0.036186094	0
Ward.D NormalizedVarInfo	55	0.036186094	0
Single NormalizedVarInfo	55	0.036186094	0
Complete NormalizedVarInfo	55	0.036186094	0
Average NormalizedVarInfo	55	0.030691906	0
Mcquitty NormalizedVarInfo	55	0.034679621	0
Median NormalizedVarInfo	55	0.035807944	0
Centroid NormalizedVarInfo	52	0.029035485	0
Ward.D2 NormalizedVarInfo	52	0.029035485	0
Ward.D Max	52	0.029035485	0
Single Max	52	0.029035485	0
Complete Max	55	0.028871751	6
Average Max	55	0.022147709	0
Mcquitty Max	55	0.029816715	0
Median Max	53	0.038300229	0
Centroid Max	55	0.038024155	3
Ward.D2 Max	53	0.038300229	0
Ward.D NormalizedMax	53	0.038300229	0
Single NormalizedMax	55	0.031336635	0
Complete NormalizedMax	55	0.037312802	0
Average NormalizedMax	55	0.036410539	0
Mcquitty NormalizedMax	52	0.035490959	0
Median NormalizedMax	52	0.035490959	0
Centroid NormalizedMax	52	0.035490959	0
Ward.D2 NormalizedMax	52	0.035490959	0
Ward.D Pearson	55	0.027138957	0
Single Pearson	55	0.03563758	0
Complete Pearson	55	0.036750103	6
Average Pearson	53	0.033583857	0
Mcquitty Pearson	55	0.033139404	3
Median Pearson	53	0.033583857	0
Centroid Pearson	53	0.033583857	0
Ward.D2 Pearson	55	0.028288773	0
Ward.D Spearman	55	0.033238554	0
Single Spearman	55	0.03497448	0
Complete Spearman	53	0.03134728	0
Average Spearman	55	0.025401263	0
Mcquitty Spearman	53	0.03134728	0
Median Spearman	53	0.03134728	0
Centroid Spearman	55	0.031521758	0
Ward.D2 Spearman	55	0.037998902	0
Ward.D Kendall	55	0.034842454	3
Single Kendall	53	0.022380431	0
Complete Kendall	55	0.021460333	3
Average Kendall	53	0.022380431	0
Mcquitty Kendall	53	0.022380431	0
Median Kendall	55	0.026309186	0
Centroid Kendall	55	0.019622355	0
Ward.D2 Kendall	55	0.018096988	0

Table 2. Centrality Measures of Dendrograms in the Complete Graph -Period II

	Degree	Closeness	Betweenness
Ward.D VarInfo	53	0.017362	0
Single VarInfo	55	0.017075	0
Complete VarInfo	53	0.017362	0
Average VarInfo	53	0.017362	0
Mcquitty VarInfo	55	0.017118	0
Median VarInfo	55	0.016621	0
Centroid VarInfo	55	0.016619	0
Ward.D2 VarInfo	52	0.029889	0
Ward.D NormalizedVarInfo	52	0.029889	0
Single NormalizedVarInfo	52	0.029889	0
Complete NormalizedVarInfo	52	0.029889	0
Average NormalizedVarInfo	55	0.023965	0
Mcquitty NormalizedVarInfo	55	0.030666	0
Median NormalizedVarInfo	55	0.030115	0
Centroid NormalizedVarInfo	52	0.025582	0
Ward.D2 NormalizedVarInfo	52	0.025582	0
Ward.D Max	52	0.025582	0
Single Max	52	0.025582	0
Complete Max	55	0.022257	0
Average Max	55	0.023851	0
Mcquitty Max	55	0.020194	0
Median Max	52	0.030415	2.25
Centroid Max	52	0.030415	2.25
Ward.D2 Max	52	0.030415	2.25
Ward.D NormalizedMax	52	0.030415	2.25
Single NormalizedMax	55	0.02421	0
Complete NormalizedMax	55	0.02935	0
Average NormalizedMax	55	0.030558	6
Mcquitty NormalizedMax	52	0.025355	0
Median NormalizedMax	52	0.025355	0
Centroid NormalizedMax	52	0.025355	0
Ward.D2 NormalizedMax	52	0.025355	0
Ward.D Pearson	55	0.022488	0
Single Pearson	55	0.029665	0
Complete Pearson	55	0.029726	6
Average Pearson	53	0.025421	0
Mcquitty Pearson	55	0.025184	3
Median Pearson	53	0.025421	0
Centroid Pearson	53	0.025421	0
Ward.D2 Pearson	55	0.022595	0
Ward.D Spearman	55	0.026503	0
Single Spearman	55	0.025504	0
Complete Spearman	53	0.025873	0
Average Spearman	55	0.019348	0
Mcquitty Spearman	53	0.025873	0
Median Spearman	53	0.025873	0
Centroid Spearman	55	0.023216	0
Ward.D2 Spearman	55	0.029771	6
Ward.D Kendall	55	0.028277	0
Single Kendall	55	0.018288	0
Complete Kendall	55	0.017318	3
Average Kendall	55	0.018288	0
Mcquitty Kendall	55	0.018288	0
Median Kendall	55	0.01911	0
Centroid Kendall	55	0.021268	0
Ward.D2 Kendall	55	0.016594	0

3.3. A Tanglegram of Dendrograms

A tanglegram is a pair of dendrograms on the same objects. It is developed to compare those two dendrograms. This method connects the studied cryptocurrencies from one dendrogram to another dendrogram. If the generated lines are all parallel, the dendrograms are the same, and if these lines intersect, it shows their divergence. The more parallel these lines are, the more similar the corresponding clusterings are, and the more intertwined these lines are, the less similar the dendrograms are to each other. As an illustration, Figure 11 shows a tanglegram to which the corresponding dendrograms are the same, and so all the lines connecting the entities are all parallel. Thus, what if we compare the best dendrogram of the period II (Normalized Max distance and Ward.D method) with the best dendrogram of the period I (Pearson's distance and Complete method)? The tanglegram from this comparison is shown in Figure 12.

Figure 12 shows that most of the connections between the two dendrograms are cut off instead of parallel. It represents the change in clusters between periods I and II. Given that the basis for dividing the research period into these two sub-periods was the COVID-19 pandemic, it can be argued that the clustering structure changed before and after this event. Decision-makers are advised to update their financial clusters based on data following the COVID-19 pandemic.

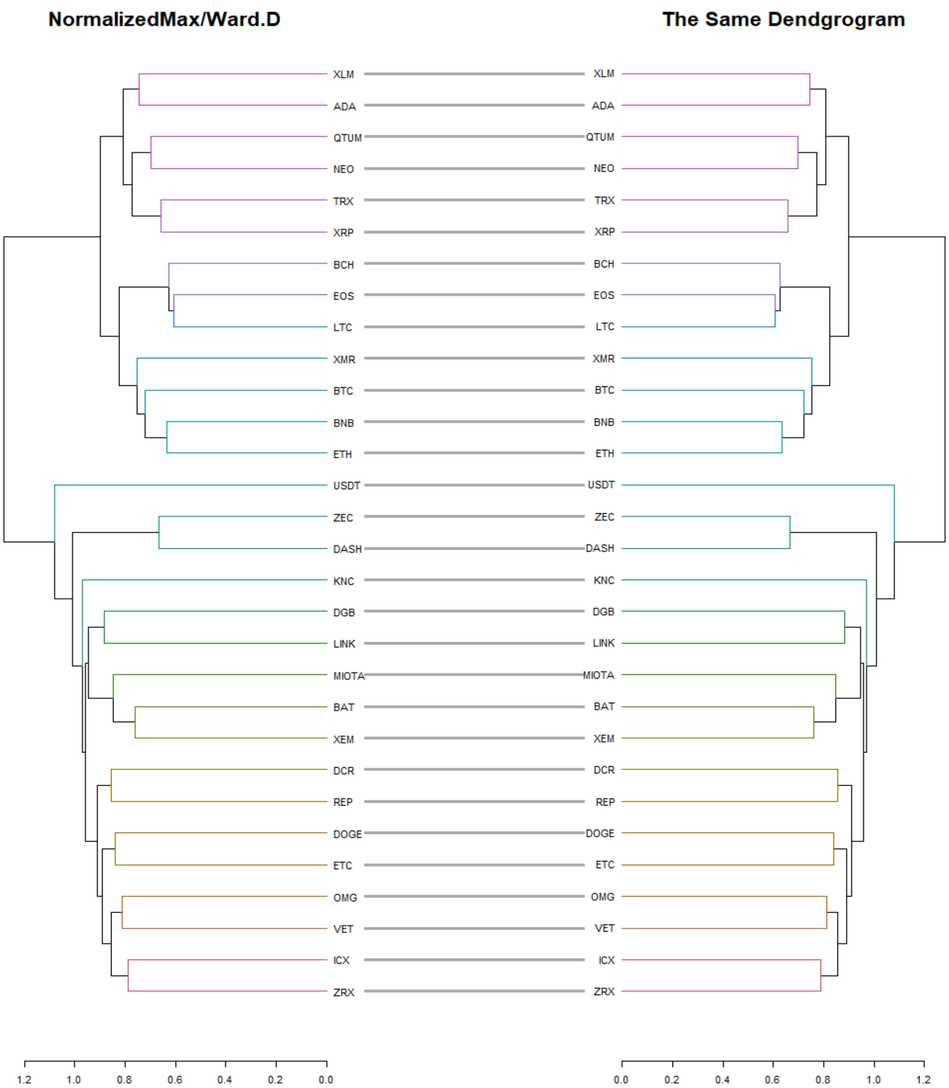

Figure 11. An Example of Tanglegram With Full Similarity

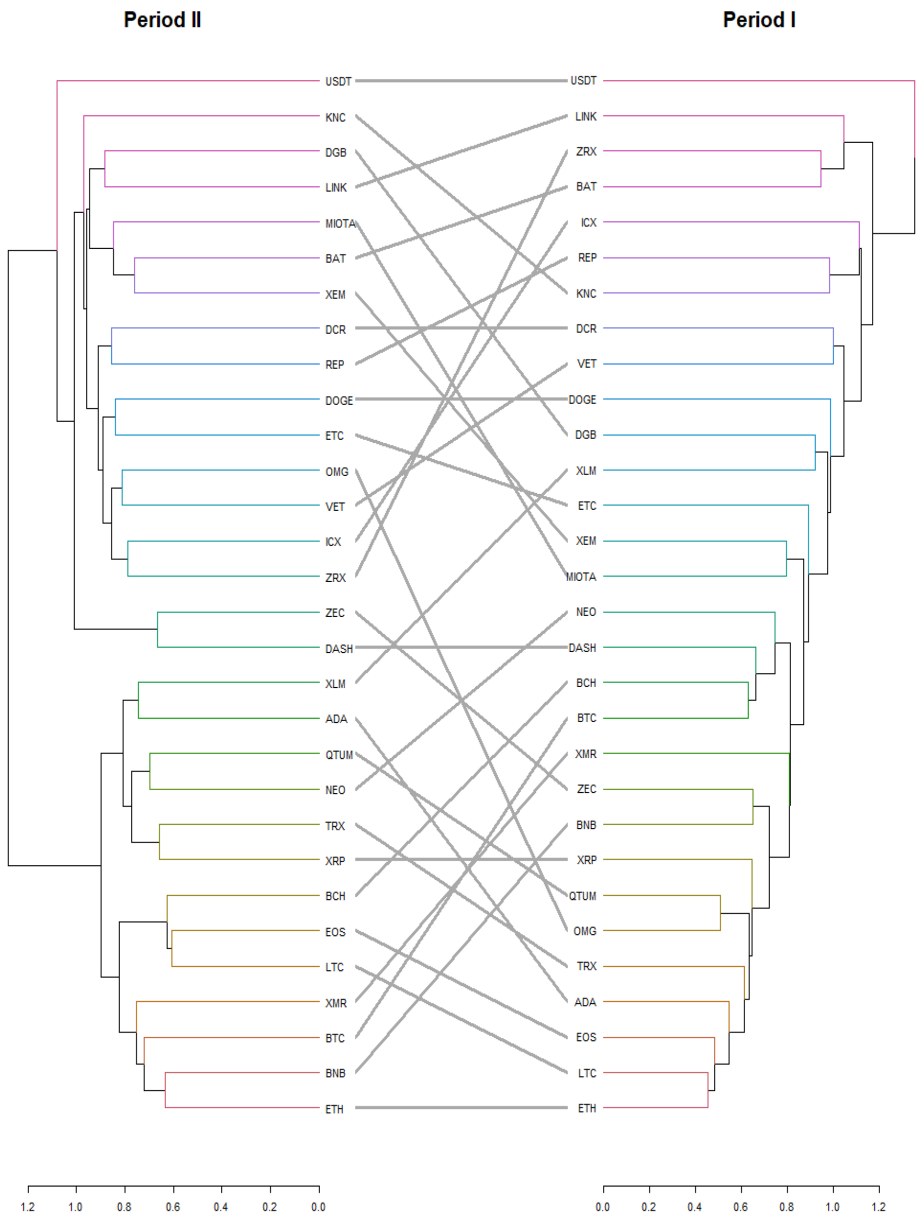

Figure 12. The Tanglegram of the Best Dendrograms: Periods I and II

4. Conclusion

This study showed in the first place that the structure of the cryptocurrency market differed before and after the COVID-19 pandemic. The reason for this and the explanation of how COVID-19 affects the structural relationships of cryptocurrencies require separate research. Some studies that have examined this issue have examined how COVID-19 affects factors such as market bubble (Montasser et al., 2021) , herd behaviors (Rubbaniy et al., 2021; Yarovaya et al., 2021), long-term market memory (Lahmiri & Bekiros, 2021), volatility, and liquidity (Corbet et al., 2021).

In this research, we studied the clustering of cryptocurrencies. The scope of the study included 30 cryptocurrencies that were most popular in financial transactions and had the highest market values. Although several studies have been conducted on cryptocurrencies, the unique feature of this research was the application of the information-based distance matrices in financial clusterings. The methodical problem in hierarchical clustering analysis is the diversity of distances as well as the existence of various methods of clustering for organizing clusters in a hierarchy. Each technique of distance computation of entities and each method for finding clusters leads to a distinct clustering (or dendrogram). In this study, we examined 56 different dendrograms of the 30 cryptocurrencies of interest, some of which were proposed with information-based distances and others with correlation-based distances. To select the best hierarchical clustering, we introduced a method called second-order hierarchical clustering.

In this method, we calculate the correlation of the dendrograms, obtain the distance of those dendrograms, and cluster them. The result is a dendrogram of dendrograms. However, this clustering can also be done in several ways, and thus, a vicious circle occurs. To stop this vicious circle, we used graph theory. In this theory, using the minimum spanning tree method, we filtered the complete graph obtained from the clusters and analyzed the resulting tree. For this purpose, we calculated the centrality indices for each vertex of this tree. Then, comparing the centrality measures of the nodes, we concluded that the most influential dendrogram corresponded to the hierarchical clustering with Normalized Max Information distance and the clustering method of Ward.D in the period II. Furthermore, the best clustering of the period I is the dendrogram obtained with Pearson’s distance and the complete method.

Clusters from the most influential dendrograms showed that some clusters were single-member, some two-member, and some multi-member. Single-member clusters indicated a considerable distance from other cryptocurrencies. This distance can have some reasons, which can be the subject of another study. In two-member clusters, there were two cryptocurrencies whose pairing was applicable in pair trading. In pair trading or statistical arbitrage, traders identify the pairs of assets that are closest to each other based on the criteria such as the correlation of return on assets or their price cointegration. They hedge their investment risk by trading these pairs in their portfolio. The hierarchical nature of this clustering allows researchers to identify sub-clusters in larger clusters. Larger clusters, however, can also be used as alternatives to an asset. For example, if a person intends to buy and sell bitcoin and for any reason is not able to trade it, he can use bitcoin alternatives. In the last step, we compared the top clusterings and showed the impact of COVID-19 on the structure of hierarchical clusterings.

As stated in the research literature, a wide range of decision makers can use this research results. First of all, cryptocurrency traders looking to select a portfolio can use the combination of the proposed clusters for diversification. In addition, members in a cluster can be good candidates for pairs trading (statistical arbitrage). Systemic risk scholars can also examine the systemic effects of a representative of each cluster on economic variables using methods such as copula modeling. In fact, this study somehow sought to explore the structural relationships of cryptocurrencies with each other. For any financial study or economic decision in which such a structure is important, the results of our research can contain effective guidelines.

References

Aggarwal, D. (2019). Do bitcoins follow a random walk model? *Research in Economics*, *73*(1), 15–22. https://doi.org/10.1016/j.rie.2019.01.002

Aghabozorgi, S., Seyed Shirkhorshidi, A., & Ying Wah, T. (2015). Time-series clustering – A decade review. *Information Systems*, *53*, 16–38. https://doi.org/10.1016/j.is.2015.04.007

Albatineh, A. N., Niewiadomska-Bugaj, M., & Mihalko, D. (2006). On similarity indices and correction for chance agreement. *Journal of Classification*, *23*(2), 301–313. https://doi.org/10.1007/s00357-006-0017-z

Amigó, J., Balogh, S., & Hernández, S. (2018). A brief review of generalized entropies. *Entropy*, *20*(11), 813-834. https://doi.org/10.3390/e20110813

Andrada-Félix, J., Fernandez-Perez, A., & Sosvilla-Rivero, S. (2020). Distant or close cousins: Connectedness between cryptocurrencies and traditional currencies volatilities. *Journal of International Financial Markets, Institutions and Money*, *67*, 101219. https://doi.org/10.1016/j.intfin.2020.101219

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Batina, L., Gierlichs, B., Prouff, E., Rivain, M., Standaert, F.-X., & Veyrat-Charvillon, N. (2011). Mutual information analysis: A comprehensive study. *Journal of Cryptology*, *24*(2), 269–291. https://doi.org/10.1007/s00145-010-9084-8

Black, F., & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, *48*(5), 28–43. https://doi.org/10.2469/faj.v48.n5.28

Boako, G., Tiwari, A. K., & Roubaud, D. (2019). Vine copula-based dependence and portfolio value-at-risk analysis of the cryptocurrency market. *International Economics*, *158*, 77–90. https://doi.org/10.1016/j.inteco.2019.03.002

Bonanno, G., Caldarelli, G., Lillo, F., Micciché, S., Vandewalle, N., & Mantegna, R. N. (2004). Networks of equities in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems*, *38*(2), 363–371. https://doi.org/10.1140/epjb/e2004-00129-6

Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T. (2016). Information theory. In T. Bossomaier, L. Barnett, M. Harré, & J. T. Lizier (Eds.), *An introduction to transfer entropy* (pp. 33–63). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-319-43222-9_3

Bouri, E., Lau, C. K. M., Lucey, B., & Roubaud, D. (2019). Trading volume and the predictability of return and volatility in the cryptocurrency market. *Finance Research Letters*, *29*, 340–346. https://doi.org/10.1016/j.frl.2018.08.015

Bouri, E., Lucey, B., & Roubaud, D. (2020). Cryptocurrencies and the downside risk in equity investments. *Finance Research Letters*, *33*, 101211. https://doi.org/10.1016/j.frl.2019.06.009

Brauneis, A., & Mestel, R. (2019). Cryptocurrency-portfolios in a mean-variance framework. *Finance Research Letters*, *28*, 259–264. https://doi.org/10.1016/j.frl.2018.05.008

Cai, R., Zhang, Z., Tung, A. K., Dai, C., & Hao, Z. (2014). A general framework of hierarchical clustering and its applications. *Information Sciences*, *272*, 29–48. https://doi.org/10.1016/j.ins.2014.02.062

Chaudhari, H., & Crane, M. (2020). Cross-correlation dynamics and community structures of cryptocurrencies. *Journal of Computational Science*, *44*, 101130. https://doi.org/10.1016/j.jocs.2020.101130

Coletti, P. (2016). Comparing minimum spanning trees of the Italian stock market using returns and volumes. *Physica a: Statistical Mechanics and Its Applications*, *463*, 246–261. https://doi.org/10.1016/j.physa.2016.07.029

Corbet, S., Hou, Y., Hu, Y., Larkin, C., Lucey, B., & Oxley, L. (2021). Cryptocurrency liquidity and volatility interrelationships during the COVID-19 pandemic. *Finance Research Letters*, 45, 102137. https://doi.org/10.1016/j.frl.2021.102137

Corbet, S., Larkin, C., & Lucey, B. (2020). The contagion effects of the COVID-19 pandemic: Evidence from gold and cryptocurrencies. *Finance Research Letters*, *35*, 101554. https://doi.org/10.1016/j.frl.2020.101554

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Wiley and Chichester. http://www.loc.gov/catdir/enhancements/fy0624/2005047799-d.html

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems* , 1695, 1-9.

Deza, M., & Deza, E. (Eds.). (2013). *Encyclopedia of distances*. Springer.

Dubitzky, W., Wolkenhauer, O., Cho, K.-H., & Yokota, H. (2013). *Encyclopedia of Systems Biology*. Springer New York. https://doi.org/10.1007/978-1-4419-9863-7

Gates, A. J., & Ahn, Y.-Y. (2019). CluSim: a python package for calculating clustering similarity. *Journal of Open Source Software*, *4*(35), 1264. https://doi.org/10.21105/joss.01264

Galili, T. (2015). Dendextend: An R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics*. 31:3718–3720. https://doi.org/10.1093/bioinformatics/btv428

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, *31*(7). https://doi.org/10.18637/jss.v031.i07

Golbeck, J. (2013). Network structure and measures. In J. Golbeck (Ed.), *Analyzing the social web* (pp. 25–44). Morgan Kaufmann/Elsevier. https://doi.org/10.1016/B978-0-12-405531-5.00003-1

Guo, X., Zhang, H., & Tian, T. (2018). Development of stock correlation networks using mutual information and financial big data. *PLOS ONE*, *13*(4), e0195941. https://doi.org/10.1371/journal.pone.0195941

Hu, B., Bi, L., & Dai, S. (2017). Information distances versus entropy metric. *Entropy*, *19*(6), 260-269. https://doi.org/10.3390/e19060260

Jang, W., Lee, J., & Chang, W. (2011). Currency crises and the evolution of foreign exchange market: Evidence from minimum spanning tree. *Physica a: Statistical Mechanics and Its Applications*, *390*(4), 707–718. https://doi.org/10.1016/j.physa.2010.10.028

Jo, S. K., Kim, M. J., Lim, K., & Kim, S. Y. (2018). Correlation analysis of the Korean stock market: Revisited to consider the influence of foreign exchange rate. *Physica a: Statistical Mechanics and Its Applications*, *491*, 852–868. https://doi.org/10.1016/j.physa.2017.09.071

Khedmati, M., & Azin, P. (2020). An online portfolio selection algorithm using clustering approaches and considering transaction costs. *Expert Systems with Applications*, *159*, 113546. https://doi.org/10.1016/j.eswa.2020.113546

Kraskov, A., & Grassberger, P. (2009). MIC: Mutual information based hierarchical clustering. In F. Emmert-Streib & M. Dehmer (Eds.), *Information theory and statistical learning* (pp. 101–123). Springer. https://doi.org/10.1007/978-0-387-84816-7_5

Kristjanpoller, W., Bouri, E., & Takaishi, T. (2020). Cryptocurrencies and equity funds: Evidence from an asymmetric multifractal analysis. *Physica a: Statistical Mechanics and Its Applications*, *545*, 123711. https://doi.org/10.1016/j.physa.2019.123711

Lahmiri, S., & Bekiros, S. (2020). The impact of COVID-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets. *Chaos, Solitons & Fractals*, *138*, 109936. https://doi.org/10.1016/j.chaos.2020.109936

Lahmiri, S., & Bekiros, S. (2021). The effect of COVID-19 on long memory in returns and volatility of cryptocurrency and stock markets. *Chaos, Solitons & Fractals*, *151*, 111221. https://doi.org/10.1016/j.chaos.2021.111221

Li, M. (2006). Information distance and its applications. In O. H. Ibarra & H.-C. Yen (Eds.), *Lecture Notes in Computer Science, Implementation and application of automata* (pp. 1–9). Springer.

Li, M., Chen, X., Li, X., Ma, B., & Vitanyi, P. M. B. (2004). The similarity metric. *IEEE Transactions on Information Theory*, *50*(12), 3250–3264. https://doi.org/10.1109/TIT.2004.838101

Liu, W. (2019). Portfolio diversification across cryptocurrencies. *Finance Research Letters*, *29*, 200–205. https://doi.org/10.1016/j.frl.2018.07.010

Liu, W, Semeyutin, A., Lau, C. K. M., & Gozgor, G. (2020). Forecasting value-at-risk of cryptocurrencies with risk metrics type models. *Research in International Business and Finance*, *54*, 101259. https://doi.org/10.1016/j.ribaf.2020.101259

Mallqui, D. C. A., & Fernandes, R. A. S. (2019). Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. *Applied Soft Computing*, *75*, 596–606. https://doi.org/10.1016/j.asoc.2018.11.038

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems*, *11*(1), 193–197. https://doi.org/10.1007/s100510050929

Meila, M. (2003). Comparing clusterings by the variation of information. In G. Goos, J. Hartmanis, & J. van Leeuwen (Eds.), *Learning theory and kernel machines* (pp. 173–187). Springer Berlin Heidelberg.

Merediz-Sola, I., & Bariviera, A. F. (2019). A bibliometric analysis of bitcoin scientific production. *Research in International Business and Finance*, *50*, 294–305. https://doi.org/10.1016/j.ribaf.2019.06.008

Meyer, P. E. (2014). *Infotheo: Information-theoretic measures*. https://CRAN.R-project.org/package=infotheo

Mnif, E., Jarboui, A., & Mouakhar, K. (2020). How the cryptocurrency market has performed during COVID 19? A multifractal analysis. *Finance Research Letters*, *36*, 101647. https://doi.org/10.1016/j.frl.2020.101647

Montasser, G. E., Charfeddine, L., & Benhamed, A. (2021). COVID-19, cryptocurrencies bubbles and digital market efficiency: Sensitivity and similarity analysis. *Finance Research Letters*, 46, 102362. https://doi.org/10.1016/j.frl.2021.102362

Okorie, D. I., & Lin, B. (2020). Crude oil price and cryptocurrencies: Evidence of volatility connectedness and hedging strategy. *Energy Economics*, *87*, 104703. https://doi.org/10.1016/j.eneco.2020.104703

Ortmann, R., Pelster, M., & Wengerek, S. T. (2020). COVID-19 and investor behavior. *Finance Research Letters*, 37, 101717. https://doi.org/10.1016/j.frl.2020.101717

Platanakis, E., & Urquhart, A. (2019). Portfolio management with cryptocurrencies: The role of estimation risk. *Economics Letters*, *177*, 76–80. https://doi.org/10.1016/j.econlet.2019.01.019

Principe, J. C. (2010). *Information theoretic learning*. Springer New York. https://doi.org/10.1007/978-1-4419-1570-2

Rasmussen, M. (Ed.). (2014). *Quantitative portfolio optimisation, asset allocation and risk management*. Palgrave Macmillan.

Rehman, M. U., & Vinh Vo, X. (2020). Cryptocurrencies and precious metals: A closer look from diversification perspective. *Resources Policy*, *66*, 101652. https://doi.org/10.1016/j.resourpol.2020.101652

Rubbaniy, G., Polyzos, S., Rizvi, S. K. A., & Tessema, A. (2021). COVID-19, lockdowns and herding towards a cryptocurrency market-specific implied volatility index. *Economics Letters*, *207*, 110017. https://doi.org/10.1016/j.econlet.2021.110017

Ryan, J. A., & Ulrich. J. M. (2020). Quantmod: Quantitative financial modelling framework. https://CRAN.R-project.org/package=quantmod

Scornavacca, C., Zickmann, F., & Huson, D. H. (2011). Tanglegrams for rooted phylogenetic trees and networks. *Bioinformatics (Oxford, England)*, *27*(13), i248-i256. https://doi.org/10.1093/bioinformatics/btr210

Sokal, R. R., & Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *TAXON*, *11*(2), 33–40. https://doi.org/10.2307/1217208

Song, J. Y., Chang, W., & Song, J. W. (2019). Cluster analysis on the structure of the cryptocurrency market via Bitcoin–Ethereum filtering. *Physica a: Statistical Mechanics and Its Applications*, *527*, 121339. https://doi.org/10.1016/j.physa.2019.121339

Song, L., Langfelder, P., & Horvath, S. (2012). Comparison of co-expression measures: Mutual information, correlation, and model based indices. BMC Bioinformatics, 13(1), 1–21. https://doi.org/10.1186/1471-2105-13-328

Steuer, R., Daub, C. O., Selbig, J., & Kurths, J. (2005). Measuring distances between variables by mutual information. In D. Baier & K.-D. Wernecke (Eds.), *Studies in classification, data analysis, and knowledge organization, innovations in classification, data science, and information systems* (pp. 81–90). Springer Berlin Heidelberg and Springer e-books.

Stosic, D., Stosic, D., Ludermir, T. B., & Stosic, T. (2018). Collective behavior of cryptocurrency price changes. *Physica a: Statistical Mechanics and Its Applications*, *507*, 499–509. https://doi.org/10.1016/j.physa.2018.05.050

Tabak, B. M., Serra, T. R., & Cajueiro, D. O. (2010). Topological properties of stock market networks: The case of Brazil. *Physica a: Statistical Mechanics and Its Applications*, *389*(16), 3240–3249. https://doi.org/10.1016/j.physa.2010.04.002

Tapscott, D., & Tapscott, A. (2016). *Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world*. Portfolio / Penguin.

Tiwari, A. K., Raheem, I. D., & Kang, S. H. (2019). Time-varying dynamic conditional correlation between stock and cryptocurrency markets using the copula-ADCC-EGARCH model. *Physica a: Statistical Mechanics and Its Applications*, *535*, 122295. https://doi.org/10.1016/j.physa.2019.122295

Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences*, *102*(30), 10421–10426. https://doi.org/10.1073/pnas.0500298102

Tumminello, T., Di Matteo, T., Aste, T., & Mantegna R. N. (2007). Correlation based networks of equity returns sampled at different time horizons. *The European Physical Journal B*, *55*(2), 209–217. https://doi.org/10.1140/epjb/e2006-00414-4

Umar, Z., & Gubareva, M. (2020). A time-frequency analysis of the impact of the Covid-19 induced panic on the volatility of currency and cryptocurrency markets. *Journal of Behavioral and Experimental Finance*, *28*, 100404. https://doi.org/10.1016/j.jbef.2020.100404

Van der Hoef, H., & Warrens, M. J. (2019). Understanding information theoretic measures for comparing clusterings. Behaviormetrika, 46(2), 353–370. https://doi.org/10.1007/s41237-018-0075-7

Von Foerster, H. (2002). Cybernetics of cybernetics. In H. von Foerster (Ed.), *Understanding understanding* (pp. 283–286). Springer. https://doi.org/10.1007/0-387-21722-3_13

Warren Liao, T. (2005). Clustering of time series data – A survey. *Pattern Recognition*, *38*(11), 1857–1874. https://doi.org/10.1016/j.patcog.2005.01.025

Warrens, M. J. (2019). Similarity measures for 2 × 2 tables. *Journal of Intelligent & Fuzzy Systems*, *36*(4), 3005–3018. https://doi.org/10.3233/JIFS-172291

Xu, Q., Zhang, Y., & Zhang, Z. (2021). Tail-risk spillovers in cryptocurrency markets. *Finance Research Letters*, *38*, 101453. https://doi.org/10.1016/j.frl.2020.101453

Yarovaya, L., Matkovskyy, R., & Jalan, A. (2021). The effects of a “black swan” event (COVID-19) on herding behavior in cryptocurrency markets. *Journal of International Financial Markets, Institutions and Money*, 75,101321. https://doi.org/10.1016/j.intfin.2021.101321

Zhang, W., Li, Y., Xiong, X., & Wang, P. (2021). Downside risk and the cross-section of cryptocurrency returns. *Journal of Banking & Finance*, *133*, 106246. https://doi.org/10.1016/j.jbankfin.2021.106246

Zięba, D., Kokoszczyński, R., & Śledziewska, K. (2019). Shock transmission in the cryptocurrency market. Is Bitcoin the most influential? *International Review of Financial Analysis*, *64*, 102–125. https://doi.org/10.1016/j.irfa.2019.04.009