ISSN: 1304-4796
Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi
Manisa Celal Bayar University Journal of Social Sciences
2025; 23 (4);549-575

Association Rules Analysis with Apriori Algorithm and Data Visualization: *Evidence from Borsa Istanbul*

Turan Kocabıyık[1, 2], Hafize Gonca Cömert[3], Meltem Karaatlı[4]

ABSTRACT

Considering that association analysis can be applied in finance, as it is in many fields, a study was conducted using stock prices. This study utilized two datasets covering the stocks included in the BIST TUM index and the BIST TUM-100 index, which were calculated within Borsa Istanbul. The data were analyzed using the Apriori algorithm and supported by data visualization techniques. The research used daily closing data of 421 companies in the BIST TUM index and 340 companies in the BIST TUM-100 index, obtained from the Finnet database. The data belongs to the period 10.01.2022/08.01.2024, and the association analysis was performed with stock closing prices according to the interestingness rule criteria. ARules () and ARulesViz () packages were used for the Apriori algorithm in the analysis performed in the R programming language, and Shiny () was used for the representation of codes and visuals. In Model 1, which is based on BIST TUM index stocks, 66 association rules were obtained, and in Model 2, which covers BIST TUM-100 index stocks, 34 association rules were obtained. The study concludes that all the parameters analyzed reveal significant and interesting rules, as statistically evidenced by Fisher's exact p-values, indicating that the companies may be significantly related.

Keywords*: Portfolio Management, Data Mining, Apriori Algorithm, Fisher's Exact Test, Interestingness Measures, Stocks.*

[1] Contact: turankocabiyik@sdu.edu.tr
[2] Assoc. Prof. Dr., Süleyman Demirel University, Faculty of Economics and Administrative Sciences, Turkey, ORCID: 0000-0003-3651-206X
[3] Ph. D. Student, Süleyman Demirel University, Institute of Social Sciences, Faculty of Economics and Administrative Sciences, Turkey, ORCID: 0009-0009-3345-8783
[4] Prof. Dr., Süleyman Demirel University, Faculty of Economics and Administrative Sciences, Turkey, ORCID: 0000-0002-7403-9587

(Submission: 29.08.2025 / Published:29.12.2025)
Doi Number: 10.18026/cbayarsos.1773973
Type of Research: Research Article

Apriori Algoritması ile Birliktelik Kuralları Analizi ve Veri Görselleştirme: *Borsa İstanbul Örneği*

ÖZET

Birliktelik analizinin birçok alanda olduğu gibi finans alanında da uygulanabileceği düşünülerek, hisse senedi fiyatları kullanılarak, bir araştırma yapılmıştır. Bu çalışmada; Borsa İstanbul bünyesinde hesaplanan BIST TUM endeksi ve BIST TUM-100 endeksinde yer alan hisse senetlerini kapsayan iki veri seti kullanılmıştır. Veriler Apriori algoritması kullanılarak analiz edilmiş ve veri görselleştirme teknikleriyle desteklenmiştir. Araştırmada Finnet veri tabanından elde edilen BIST TUM endeksinde yer alan 421 şirket ile BIST TUM-100 endeksinde yer alan 340 şirketin günlük kapanış verileri kullanılmıştır. Veriler 10.01.2022/08.01.2024 dönemine ait olup, birliktelik analizi, ilginçlik ölçütlerine göre hisse senedi kapanış fiyatları ile gerçekleştirilmiştir. R programlama dilinde gerçekleştirilen analizlerde Apriori algoritması için ARules () ve ARulesViz () paketleri, kodların ve görsellerin gösterimi için ise Shiny () kullanılmıştır. BIST TUM endeksi hisse senetlerini kapsayan Model 1'de 66, BIST TUM-100 endeksi hisse senetlerini kapsayan Model 2'de ise 34 birliktelik kuralı elde edilmiştir. Çalışma, analiz edilen tüm parametrelerin, Fisher'in kesin p-değerleri ile istatistiksel olarak kanıtlandığını, şirketlerin önemli ölçüde ilişkili olabileceğini gösteren, önemli ve ilginç kuralları ortaya koyduğu sonucuna varmaktadır.

Anahtar Kelimeler: *Portföy Yönetimi, Veri Madenciliği, Apriori Algoritması, Fisher's Exact Test, İlginçlik Ölçütleri, Hisse Senetleri.*

1. INTRODUCTION

Data mining can be applied to a wide variety of fields. In particular, marketing, finance, electronic commerce, health, and social media can be counted among them (Özkan, 2016). In this study, an application on finance was carried out and the associations of the daily closing prices of the stocks included in two indices selected from the Borsa Istanbul (BIST) equity market were investigated. For this purpose, the associations of the stocks were analyzed and visualized by using historical data.

Borsa Istanbul is the market where a significant portion of capital market activities in Turkey take place. The equity market is one of the four markets of the stock exchange. Indices are calculated to follow the decreases and increases in the equity market. BIST TUM and BIST TUM-100 indices are also among these indices. BIST TUM index is the index that includes all the stocks in the market. BIST TUM-100 is the index that consists of the other stocks in the BIST TUM index except those in the BIST-100 index (borsaistanbul.com).

There is rapid change, development, and innovation in the world, information is produced rapidly, the information produced is transformed into technology, and the information created in the same way becomes obsolete and is replaced by new information and new fields of science. These dizzying changes affect the learning-teaching processes, and the methods and techniques of accessing information, learning it, transferring it to life, and using it in solving a problem are changing. These developments have led to the emergence of data mining since the 1990s and the way of discovering information by analyzing the existing database (Cemaloğlu & Duykuluoğlu, 2020).

There is no single, fixed definition for "What is data mining?", but many different definitions have been made. Although data mining is used in the sense of mining information from data, information extraction, data/pattern analysis, data archaeology, and data scanning, it is the process of mining a large amount of data to reveal a small nugget, that is, information, from hidden patterns (Han et al., 2012).

Data mining is the application of special algorithms to extract patterns from data (Fayyad et al., 1996). As a short and general definition, data mining means mining on data (Erden, 2021). Data mining is a process. This process starts from cleaning the data; integrating, reducing, transforming, applying data mining methods, and evaluating the results (Özkan, 2016). Data mining methods are generally fed by two main roots—statistics and artificial intelligence, and machine learning, an extension of artificial intelligence. In addition to these disciplines, database management systems, linguistics, visualization, and geographic information systems also play an important role in data mining methods (Akpınar, 2017).

Many methods and algorithms have been developed in data mining, and many of these methods are statistically based (Özkan, 2016). Data mining models are grouped under three main headings. These are Classification, Clustering, and Pattern Mining-Association Rules. Data mining methods in which two or more simultaneous or different occurrences are analyzed within a set of elements and association rules are developed are called pattern mining. Association rules are used to define relationships that occur simultaneously (Akpınar, 2017). The Apriori algorithm is one of the most widely used data mining methods designed to identify frequently co-occurring itemsets within large-scale databases and to generate meaningful association rules that reveal relationships among these itemsets (El Mahjouby et al., 2024). Based on Boolean association rules, this algorithm first identifies frequent itemsets and then scans the entire database to uncover relationships among them, thereby supporting decision-making processes (Xie, 2021).

As for the financial importance of this research, which can be considered interdisciplinary; it is to provide information on which instruments stock investors can systematically turn to when making investments. The stocks included in the obtained association rules should not be used in portfolio diversification to reduce risk. Because it is against the nature of the business to use investment instruments that rise and fall in the same periods while diversifying the portfolio. In addition, if the decline or rise in one of the stocks whose movements are determined associated has not yet been experienced in the other stock, it gives the investor time to buy or sell. This may create a situation that will give investors an advantage in their investment decisions.

This study consists of four sections. The first section of the study includes brief information about data mining and Borsa Istanbul, while the second section includes a literature summary. The third section is where data sets and method algorithms are discussed and analysis is performed. The fourth section, the last section, includes the analysis results and findings.

2. LITERATURE REVIEW

In this section, the literature on association analysis within the scope of data mining is presented. Firstly, the applications of association analysis to the equity and security markets are presented, followed by studies in other areas. Information on the literature is listed from recent to old. Some of the studies in the equity and security markets where association analysis is used as a method are given in Table 1.

Table 1. Literature on Equity and Security Markets

Author(s)	Year	Algorithm	Program	Application
Kocabıyık et al.	2024	FP-Growth	WEKA	Clustering and association analysis of 71 cryptocurrencies with data for the period 2021- 2022
Liu	2024	Apriori	Python	Association analysis of Shanghai Stock Exchange financial sector companies with Apriori algorithm
Pekdemir et al.	2023	Apriori		Analyzing the social marketing activities of 12 banking companies listed on Borsa Istanbul in Turkey
Kartal et al.	2022	Apriori	R-Studio	Association analysis of Brazil (BVSP), USA (S&P500, DJI, IXIC), France (FCHI), Germany (GDAXI), UK (FTSE), China (SSEC), Hong Kong (HSI), Japan (N225) and Turkey (XU100) stock exchanges
Ataman et al.	2022	Apriori, CART	WEKA	Classification and association analysis with monthly closing prices of BIST-30 and BIST-100 indices
Karaatlı et al.	2021	FP-Growth	WEKA	Association analysis of stocks in the BIST-30 index and various investment instruments
Kocabıyık et al.	2021	FP-Growth	WEKA	Association analysis of stocks in Borsa Istanbul BIST 30 index with data of 1601 trading days
Ünsal	2020	Apriori, K-means	Python, WEKA	Association analysis with 2019 data of 408 stocks traded in the Borsa Istanbul equity market

Source: Authors' Representation

Examples of studies where association analysis is used outside the field of finance are presented in Table 2.

Table 2. Literature Review Related to Non-Finance Fields

Author(s)	Year	Algorithm	Program	Application
Azis et al.	2024	Apriori	Tanagra	Delta Computer Cijantung company stock items Apriori analysis
Çelik et al.	2023	Apriori	MS SQL	Product analysis with Apriori in businesses that engage in storage activities
Oğur et al.	2023	Apriori	Python	Association analysis on customer behavior in the retail industry
İnce et al.	2022	Apriori	Tanagra	Classification and association analysis of 102 holdings operating in Turkey
Abidin et al.	2022	Apriori	Python	Association analysis with sales data from Prima Motor Shop
Bayram et al.	2021	Apriori	SPSS	Association analysis of 960 bank branches in 81 provinces of Turkey
Karasu et al.	2020	Apriori	WEKA	Association analysis in a fast-moving consumer goods company
Asur et al.	2020	Apriori	WEKA	Association analysis of 9 different landscape types in Turkey

Source: Authors' Representation

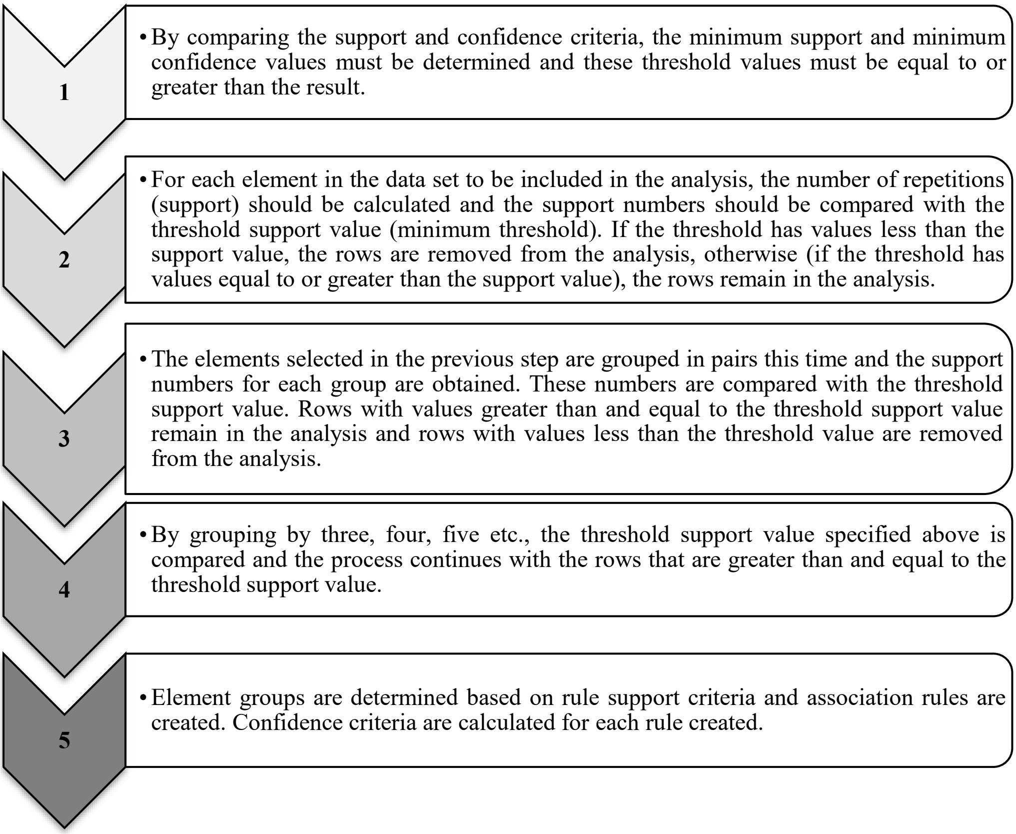
3. METHODOLOGY AND ANALYSIS

This section includes the Apriori algorithm, which is the method of the study, and the analyses. First, the parameters and functions of the Apriori algorithm are mentioned. This section, which also includes the data sets of the applications in the study, is the section where the Apriori analysis is performed and the findings obtained in the analysis are included.

3.1. Apriori Algorithm

The method that tries to reveal which events can co-occur by examining the relationship between the records in the database is called "Association Rules Analysis". Apriori, Eclat, and FP-Growth are some of the association rule algorithms (Özkan, 2016). Apriori is a groundbreaking algorithm proposed by Agrawal and Srikant in 1994 for frequency pattern mining and finding association rules (Agrawal & Srikant, 1994). The stages of the Apriori algorithm, which is most commonly used in association analysis, are presented in Figure 1 (Özkan, 2016).

Figure 1. Apriori Algorithm



Source: Özkan, Y., 2016. *Veri Madenciliği Yöntemleri Bilgisayar Bilimleri ve Bilgisayar Mühendisliği*, İstanbul, p.219.

3.2. Parameters (Interest Rule Criteria)

Apriori is an algorithm that follows a straightforward level-wise approach whose main function is calculating the frequencies of the element sequences that form the basis of association rules. The algorithm is called "Apriori" because the information obtained at the previous level is used (Akpınar, 2017). A rule in the algorithm consists of a proposition on the left (antecedent or condition) and a proposition on the right (consequence). In the rule, if the left side (antecedent (s)) is true, the right side (consequent (s)) is also true. The probability P is the simple conditional probability that the right side is also true, given that the left side is true (Hand et al., 2001). The rules in the Apriori algorithm, which are suitable for small databases, are discrete in nature and are well-suited for modeling categorical variables (Oğuzlar, 2004).

The most commonly used parameters of the Apriori algorithm are support, confidence, and lift. In addition to these basic criteria, there are other parameters (interesting rule criteria) that reveal the interestingness of association rules, and these are given in Figure 2 (Aydemir, 2019; Hahsler, 2015). All of the parameters (interesting rule criteria) are important criteria used in data analysis. They are used to understand various relationships and patterns in the data. Each of these criteria allows the analysis and evaluation of the interestingness and significance of the rules from a different perspective. While each criterion can be used in different situations depending on the analysis purposes and the characteristics of the data, whether a particular criterion is the most important usually varies according to the analysis context. Association rules are in the form of X→Y. It means that the elements in the data set that satisfy the rule X in the rule also satisfy the conditions of Y. To create association rules, the minimum support value must first be determined in advance. As this value approaches 0, the number of associations increases, and when it approaches 1, the number of associations decreases. The minimum confidence value must also be determined, and this value is generally greater than 0.50. The larger the support and confidence parameters, the stronger the association rules are, but it also causes some rules to be overlooked (Uslu, 2018; Teker and Konuşkan, 2022). Otherwise, being too small causes a departure from the desired pattern. Another important parameter is the lift value, and if this value is equal to 1, there is no relationship between the elements (it is not interesting), if it is less than 1, it indicates that X and Y are not possible. If the lift value is greater than 1, it indicates that if X occurs, there is a high probability that Y will occur, which is desired (Altunkaynak, 2022).

Functions are used after the association rules are determined according to the parameters. According to the is.redundant () function, if a rule is redundant, that rule is not significant on its own. The is.maximal () and is.significant () functions are used to determine the properties of the rules. In other words, if a rule is maximal, the rule is interesting and significant; if significant, the rule is significant and important. It is possible to evaluate the properties, importance, and comprehensiveness of the rules with these functions (Hahsler et al., 2023). Fisher's exact test, proposed by Ronald A. Fisher, tests the relationship between two categorical variables and is used to understand whether a rule is random or a real pattern. In other words, it is a statistical test that evaluates the relationship between two events (LHS and RHS) and gives reliable results, especially in small data sets. Fisher's exact test p-value is a critical criterion used to determine whether two events are independent of each other and to determine significance. If the P-value is generally below 0.05 (fisher.p.values < 0.05), the rule is considered statistically significant and there is a significant relationship between the events (rules). The lower the P-value, the more significant the rule is (Hahsler and Hornik, 2007; Hahsler, 2015).

Figure 2. Apriori Parameters

Apriori Parameters

- **Support:** Indicates how often an itemset occurs in the dataset.
- **Confidence:** It shows the probability that the right side (rhs) of a rule will occur if the left side (lhs) occurs.
- **Coverage:** It shows how common the left side (lhs) of the rule is in the dataset.
- **Lift:** It shows how strong the relationship between X and Y is compared to the independence of the rule.
- **Conviction**: Shows how often the left side of a rule does not occur given the absence of the right side.
- **Leverage**: Indicates how much more frequently the rules occur together than when they occur independently.
- **Jaccard**: It is the ratio of the probability of two elements appearing together in a rule to the probability of appearing alone.
- **Phi:** In a rule, it is the correlation coefficient that measures the linear relationship between two elements.
- **Gini**: Measures the difference between the probability of an event occurring and the probability of it not occurring, that is, the purity and uncertainty of the rule.
- **Cosine**: Normalizes the probability of both sides of a rule being seen together, relative to their individual probabilities of being seen.
- **Odds Ratio:** A coefficient that measures the strength of the relationship between two elements in a rule.
- **Cohen Kappa**: Evaluates how compatible the rules created by two different algorithms are.
- **Kulczynski**: Measures the proportion of two elements that appear together in a rule relative to the proportion of each element that appears separately.
- **Lambda** (λ)**:** It measures the relationship between categorical variables and shows how well one categorical variable predicts another.
- **Fisher.p.values:** Evaluates the relationship between two elements in a rule with Fisher's exact test.

Source: Authors' Representation

The formula and range values of the functions and parameters used in the association rule analysis are given in Table 3 (Hahsler, 2015; Hahsler et al., 2023).

Table 3. Functions and Parameters

Function	Explanation	Scope
is.redundant ()	It evaluates the redundancy of the rules created. If a rule is redundant, that rule is not meaningful on its own.	Rules
is.significant ()	Evaluates the statistical significance of a rule. If the rule is significant, that rule indicates a significant relationship.	Rules, α, p.values
is.maximal ()	Evaluates whether a rule is maximal. If a rule is maximal, it is generally more meaningful and interesting.	Itemsets, Rules, ItemMatrix

Criterion	Formula	Range
Support	Support(X → Y) = Count(X ∪ Y) / N	[0, 1]
Confidence	Confidence(X → Y) = Support(X ∪ Y) / Support(X)	[0, 1]
Coverage	Coverage(X → Y) = Support(X) = P(X)	[0, 1]
Lift	Lift(X → Y) = Support(X ∩ Y) / (Support(X) × Support(Y))	[0,∞]

Criterion	Formula	Range
Conviction	(1 - Support(Y)) / (1 - Confidence(X → Y))	[0,∞]
Leverage	Support(X ∪ Y) - (Support(X) × Support(Y))	[-1, 1]
Jaccard	Support(X ∪ Y) / (Support(X) + Support(Y) - Support(X ∪ Y))	[0, 1]
Phi	(P(X ∩ Y) - P(X)P(Y)) / sqrt(P(X)(1-P(X)P(Y)(1-P(Y))	[-1, 1]
Gini	$P(X)[P(Y\|X)^2 + P(\bar{Y}\|X)^2] + P(\bar{X})[P(Y\|\bar{X})^2 + P(\bar{Y}\|\bar{X})^2] - P(Y)^2 - P(\bar{Y})^2$	[0, 1]
Cosine	Support(X ∪ Y) / sqrt(Support(X) × Support(Y))	[0, 1]
Odds Ratio	(Confidence(X → Y) / (1 - Confidence(X → Y))) / (Confidence(¬X → Y) / (1 - Confidence(¬X → Y)))	[0,∞]
C.Kappa	$(P(X \cap Y) + P(\bar{X} \cap \bar{Y}) - P(X)P(Y) - P(\bar{X})P(\bar{Y})) / (1 - P(X)P(Y) - P(\bar{X})P(\bar{Y}))$	[-1, 1]
Kulczynski	1/2 × ((Support(X ∪ Y) / Support(X)) + (Support(X ∪ Y) / Support(Y)))	[0, 1]
Lambda	$\sum x \in X(max_{x \in Y}$ (P(X∩Y)) - $max_{x \in Y}$ (P(Y))) / (n - $max_{x \in Y}$ (P(Y)))	[0, 1]
Fisher's Exact Test	p-value = P($C_{XY} \geq n_{XY}$)	p.values<0.05

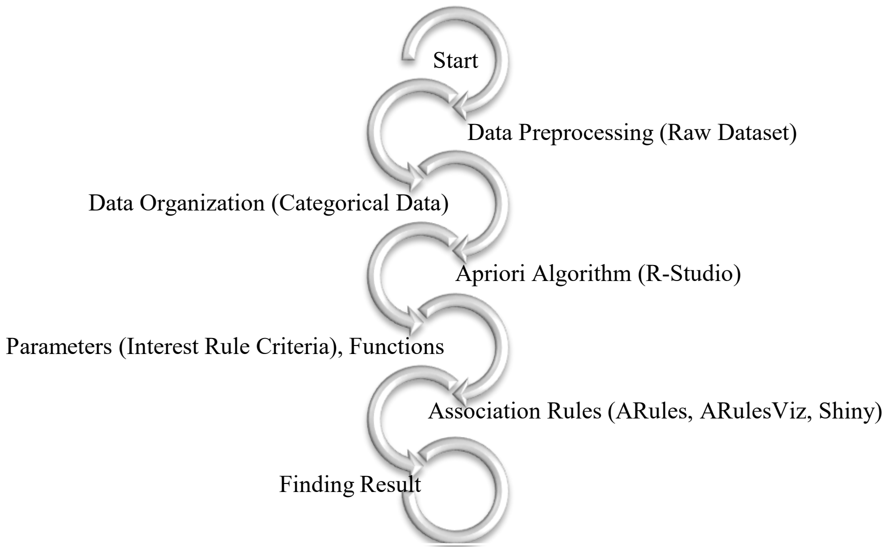
Source: Authors' Representation

3.3. Datasets

The study included two separate applications and the flow chart of the applications is presented in Figure 3. The closing prices of the stocks covering the period 10.01.2022/08.01.2024 of two groups of data sets, Application 1 (BIST TUM) and Application 2 (BIST TUM - BIST 100), were taken into account. Most studies on stocks focus on basic indices. Stocks included in the BIST 30, BIST 100, BIST FINANCIAL, and BIST SINAI indices are frequently the subject of studies. In this study, BIST TUM stocks were first included in the scope of the research. In this way, Application 1 was carried out. Then, to see the price associations of other stocks without the effect of stocks included in the BIST 100 index, which has a significant place in terms of transaction volume and includes large companies, BIST 100 stocks were removed from the BIST TUM index, and Application 2 was carried out.

The data sets included in the analysis were obtained from the Finnet database: the closing prices of 421 stocks in Application 1 and 340 stocks in Application 2. Although there are 521 stocks in the Borsa Istanbul stock market, there is missing data for 95 stocks between the dates considered for the analysis. The missing data is because these stocks are new public offerings. In addition, since the data of stocks with codes ISATR, ISBTR, KRDMA, KRDMB, and UMPAS remained the same for a long time, these stocks could not be included in the analysis. Similarly, although there are 421 stocks included in the scope of Application 2, 340 stocks were analyzed due to missing data. The only limitation of the study is that privileged stocks that have recently been offered to the public, therefore have missing data and are traded very little could not be included in the application.

Figure 3. Application Flowchart



Source: Framework Developed by the Authors for the Implementation Methodology

3.4. Apriori Analysis

For Application 1 and Application 2, the raw data sets of closing stock prices were converted to categorical (nominal) data sets. This conversion process was performed by using the “IF” function in MS Excel, taking into account the closing stock prices for both applications. If the closing stock prices increased the day before, i.e. positively (↑), it took the value 1, and if they remained the same or decreased, i.e. negatively (↓), it took the value 0. A cross-sectional example from the categorical data set of the AVOD variable in Application 1 and Application 2 is given in Table 4.

Table 4. Dataset of AVOD Variable.

	Start/End	Closing Price	Nominal Variable
Day	**Date**	**AVOD**	**AVOD**
1	08.01.2024	3.21	1
2	05.01.2024	3.18	1
3	04.01.2024	3.14	1
…	…	…	…
498	12.01.2022	2.33	0
499	11.01.2022	2.33	0
500	10.01.2022	2.36	1

Source: Finnet Database (2024)

After the preliminary preparation of the data sets, categorical data sets were prepared as MS Excel files. These data sets were then converted to a file format with the extension “.csv”. For Application 1, the data set with the extension “.csv”, belonging to 500 transaction days and 421 variables, and for Application 2, the data set with the extension. csv, belonging to 500 transaction days and 340 variables, were transferred to the R-Studio program. Association analysis was applied to these data sets using the Apriori algorithm.

To obtain significant and meaningful association rules, the is.redundant (), is.maximal () and is.significant () functions and all interestingness measures given in Table 3 were used. The analysis was performed with the ARules (1.7.7) package in the R-Studio (4.3.3) program. For the display of codes and visuals, the ARulesViz (1.5.4) and Shiny (1.8.1.1) packages were used. The Apriori algorithm in the ARules package is an algorithm that provides the extraction of association rules of frequent itemsets, maximum frequent itemsets, and near frequent itemsets. ARulesViz provides visualization techniques for association rules and itemsets obtained with the “ARules” package. Shiny is a rich package in terms of visualization techniques that facilitates the creation of interactive web applications directly from R.

3.5. Findings: Apriori Parameters and Association Rules

The minimum support and confidence coefficient parameters of the Apriori algorithm for Application 1 and Application 2 were determined as 40% and 80%, respectively, in this study.

3.5.1. Application 1

Considering the functions and parameters, in the Application 1 analysis covering the BIST TUM index, 66 association rules were formed, 48 of which were 2-variable and 18 of which were 3-variable, from 500 observation values of 421 stocks. According to the "is.redundant ()" function, these rules are "False", meaning that all the rules are not "redundant" and are meaningful on their own. According to the "is.significant ()" function, all the rules are "True", meaning that these rules represent a truly distinct relationship in the non-randomly generated data set. According to the "is.maximal ()" function, 55 rules are "True", meaning that these rules are more interesting and meaningful. 11 rules are "False", meaning that these rules are covered by another rule. The Application 1 analysis results are satisfactory and the rules represent a meaningful, interesting, and specific relationship according to the function and parameter results.

Among all rules, Rule 16 shows the most statistically significant association, with the lowest Fisher's p-value (1.05E-65). Rule 16 covers the associations {HALKB=0} => {VAKBN=0}. In fact, each criterion for this rule shows a significant and strong relationship. The association rule {HALKB=0} => {VAKBN=0} indicates that when the {HALKB} stock decreases, the {VAKBN} stock also tends to decrease or remain stable, with a support level of 0.442 and a confidence level of 0.874.

When considering association rules, the lift value should also be considered in addition to the high support and confidence parameters. Because, strong relationships are rules where the lift value is above 1, and the lift value of Rule 16 is 1.699. Accordingly, it can be said that there is a strong relationship between {HALKB} and {VAKBN}. Other interesting criteria also support that this relationship is strong and significant. All association rules can be evaluated in this way and the importance ranking can be done separately according to each of the selected parameters. A detailed evaluation of Rule 16 according to function and parameters is given in Table 5.

Table 5. Function and Parameter Evaluation According to Rule 16

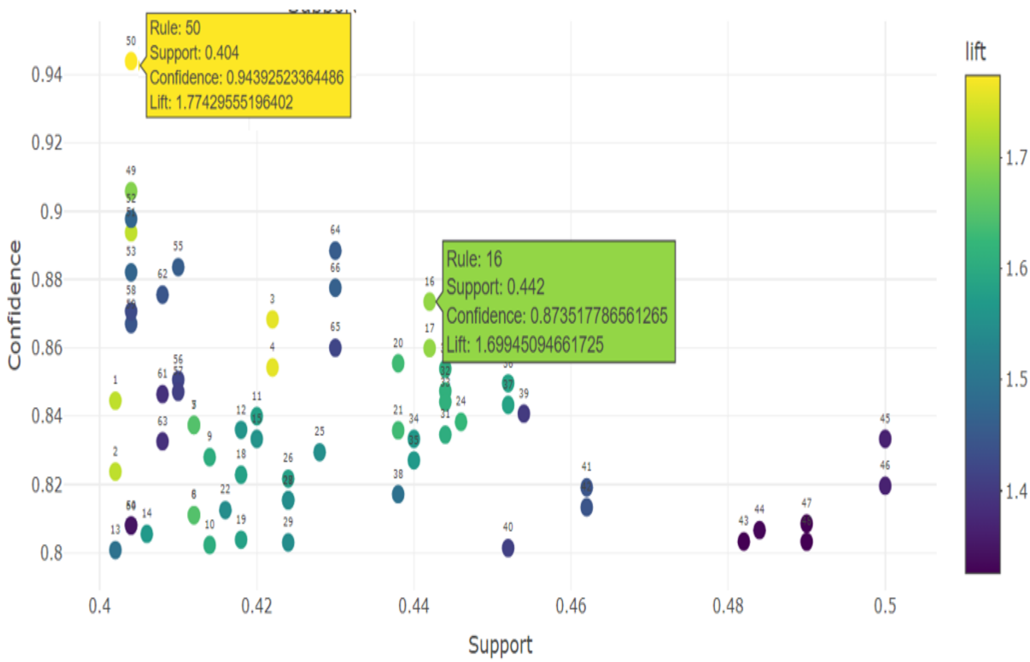
Function, Parameter	Valid Values	Value	Explanation
Redundant	(T/F)	False	The rule is not redundant, it is meaningful on its own.
Significant	(T/F)	True	The rule is meaningful, it represents a relationship.
Maximal	(T/F)	True	The rule is among the maximal rules, interesting and meaningful.
Support	[0,1]	0.442	The rule occurs in 44.2% of all transactions.
Confidence	[0,1]	0.874	VAKBN is also present in 87.4% of the transactions involving HALKB.
Coverage	[0,1]	0.506	The proportion of transactions in which HALKB and VAKBN move together constitutes 50.6% of all transactions.
Lift	[0,∞]	1.699	The probability of VAKBN exhibiting a decrease or remaining stable is 169.9% higher than the probability of HALKB independently exhibiting a decrease or remaining stable.
Conviction	[0,∞]	3.842	It shows that when HALKB=0, the probability of VAKBN=0 is low. That is, there is a strong relationship between these two events.
Leverage	[1,1]	0.182	The value of 18.2% indicates that this rule is not random and has a significant relationship in the data set.
Jaccard	[0,1]	0.765	It is the intersection and combination ratio that shows that the probability of HALKB=0 and VAKBN=0 events occurring together is quite high.
Phi	[1,1]	0.728	It shows that there is a positive and high correlation between the HALKB=0 and VAKBN=0 events.
Gini	[0,1]	0.265	A value of 26.5% indicates that the relationship in the rule is significant and that this relationship follows a certain order in the data set.
Cosine	[0,1]	0.867	It shows that there is a high similarity between HALKB=0 and VAKBN=0.
OddsRatio	[0,∞]	40.478	There is a very strong relationship between the events HALKB=0 and VAKBN=0.
CohenKappa	[1,1]	0.728	It shows that there is a good agreement between the events HALKB=0 and VAKBN=0.
Kulczynski	[0,1]	0.867	The events HALKB=0 and VAKBN=0 have a strong relationship.
Lambda	[0,1]	0.720	It is the value that measures the relational strength showing that there is a significant relationship between the events HALKB=0 and VAKBN=0.
Fisher p.values	p<0.05	1.05E-65	The Fisher p-value indicates that the rule is highly statistically significant.

Source: Authors' Representation

In Application 1 and Application 2, scatterplot, clustered heatmap matrix plot, and network diagram graph were used to better understand the association rules, and the mentioned graphs were obtained with Shiny.

The distribution graph of the 66 rules formed in Application 1 according to their support, confidence, and lift values is given in Figure 4. The balloons at the intersections of the support and confidence parameters in the graph indicate the interestingness (difference) of the rules, that is, their lift values, and these balloons are colored according to the color scale of the lift values. The yellow balloon indicates the rule with the highest lift value, while the dark blue color indicates the rule with the lowest lift value. The rule with the highest lift value is Rule 50, and the rule that is statistically highly significant according to Fisher's exact test p-value is Rule 16.

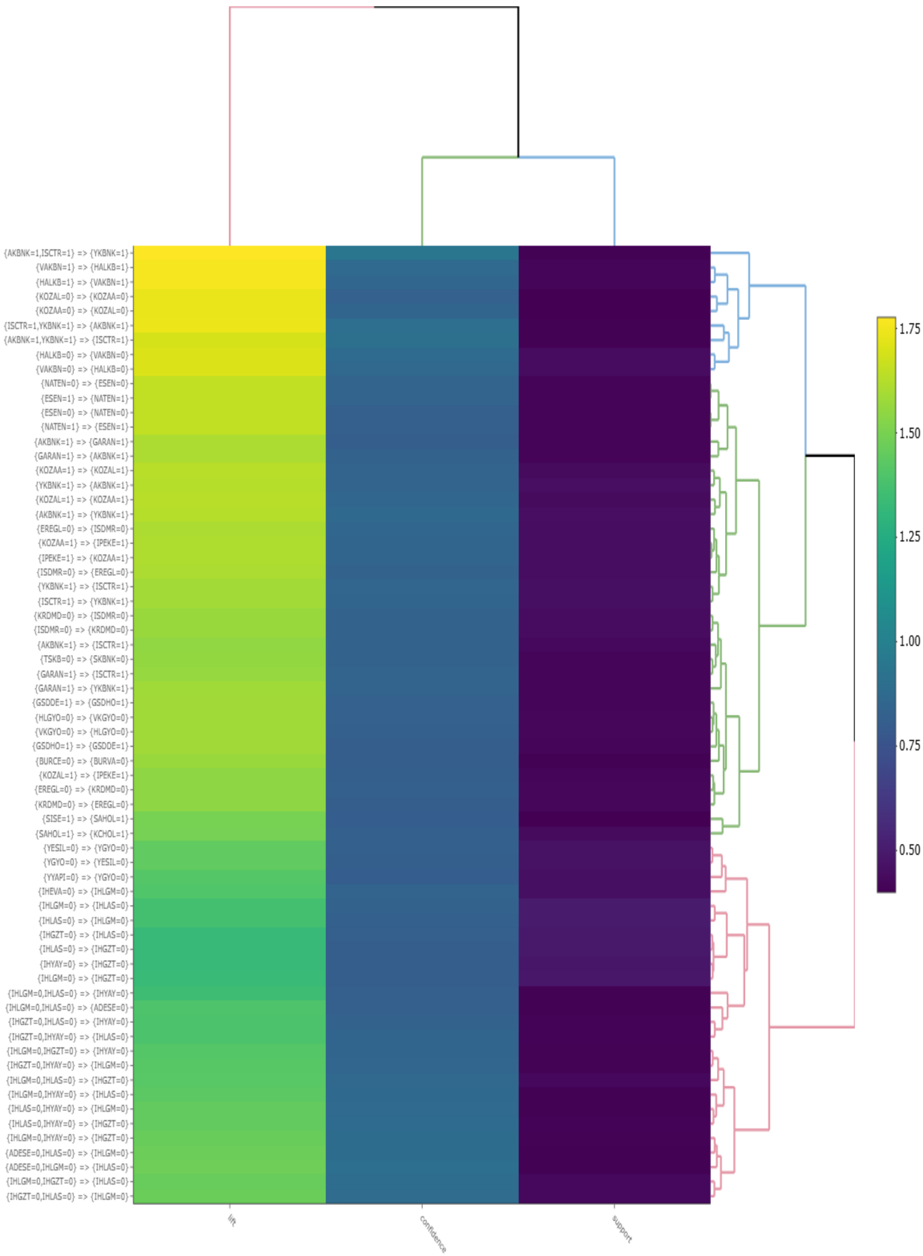
Figure 4. Scatter Plot (Application 1)



The grouped heatmap matrix graph for Application 1 is given in Figure 5. This graph is usually a matrix presented as heatmaps that show data points in groups or clusters and are used to visualize similarities and relationships in the dataset.

The rows and columns of the heatmap, which is a color-coded table, are usually ordered according to hierarchical clustering trees. Lift values are grouped according to support and confidence values in the rows according to their importance and are colored accordingly. There are 3 clusters in the grouped heatmap matrix graph that show how the rules are distributed in the dataset.

Figure 5. Grouped Heatmap Matrix Chart (Application 1)



Accordingly, the rules with a lift value of 1.5 and above, shown in blue in the dendrogram, are Rules 50, 3, 4, 2, 1, 51, 49, 16, and 17, respectively.

It can be said that there is a strong relationship between the stocks for the 9 rules in this blue cluster. {AKBNK}, {ISCTR}, and {YKBNK} stocks have strong and important relationships in the direction of increase, {KOZAA} and {KOZAL} stocks have strong and important relationships in the direction of decrease, and {VAKBN} and {HALKB} stocks have strong and important relationships in both the direction of decline and increase. In the green cluster group, where there are 32 rules, there is a rare association of companies that can be predicted correctly. There are frequent associations in the data set of 25 rules in the red cluster. In short, the grouped heatmap matrix chart allows the determination and understanding of important relationships and patterns in the evaluation of the rules together.

The network diagram that illustrates which shares the association rules of Application 1 focus on is presented in Figures 6a and 6b, where Figure 6a shows a Circular Network Diagram of Association Rules and Figure 6b displays a Force-Directed Network Diagram of Association Rules.

Figure 6a. Circular Network Diagram (Application 1)

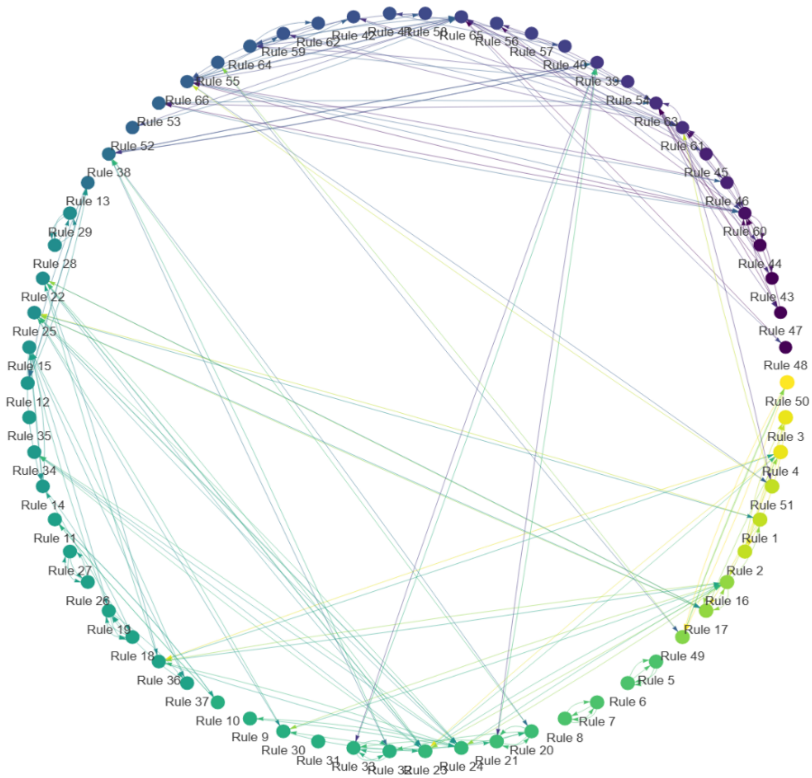
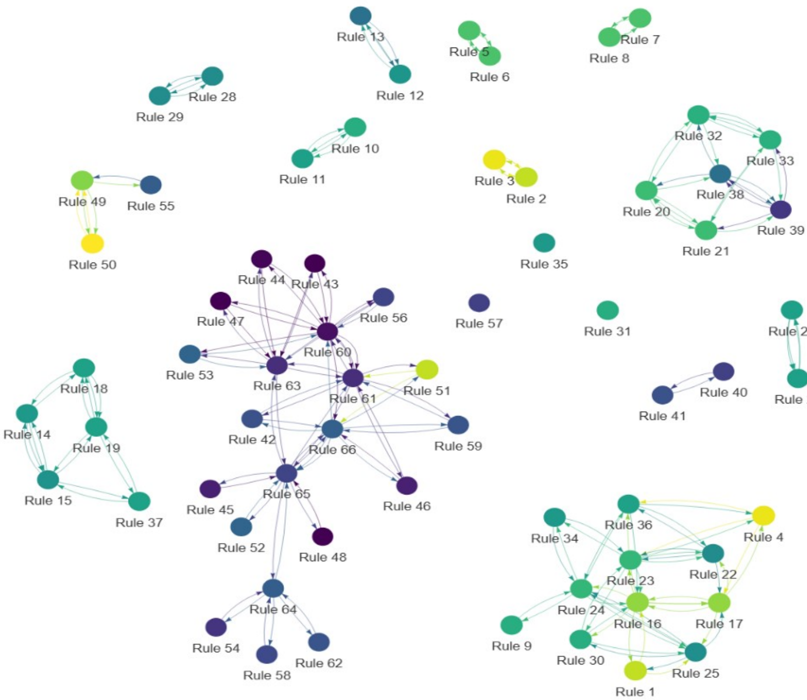


Figure 6b. Force-Directed Network Diagram (Application 1)



The network diagram graphics visualizes the association rules with nodes (balloons) and connections (edges). Here, each node represents a rule and the connections between the nodes show the relationships between the rules. In addition, the color of the node changes according to the lift values. There is a color scale from yellow to blue in the diagram, yellow nodes indicate rules with high lift values, while green indicates rules with medium, and blue indicates rules with the lowest lift values.

Rule 50 is shown in yellow because it has the highest lift value. Rule 16 is the most important rule according to the p-value, but is shown as a light green node in the graph according to its lift value. Rule 50 is connected to Rule 49. Rule 16 is connected to Rules 1, 17, 23, 24, 30, and 36. These connections indicate that these rules share a relationship or interaction between the elements in the data set. In other words, they may indicate frequently repeated relationships between certain groups of elements in the data set, rules with a common LHS or RHS, or rules with high lift and confidence values. The fact that a rule creates more than one connection is an indication that the rule has a strong and widespread relationship in the data set, as well as an important part of the rule's overall model. Rules with a high number of connections indicate that the rule has a strong relationship with other elements in the data set if they have high support and confidence values; if they have a high lift value, these relationships are not coincidental and are meaningful. In short, rules with many connections can often indicate important relationships and patterns in the data set, and examining these rules in detail can make significant contributions to the data analysis and modeling process. Therefore, these rules and their connections should be evaluated and analyzed together. When Rule 50 and its connection

Rule 49 are considered; In Rule 50, {AKBNK=1, ISCTR=1} stocks increased together at 0.404 support level and 0.944 confidence level in 202 trading days, and {YKBNK=1} stocks increased together at 1.774 lift level. In Rule 49, {AKBNK=1, YKBNK=1} stocks increased together at 0.404 support level and 0.906 confidence level in 202 trading days, and {ISCTR=1} stocks increased at 1.690 lift level. The association of these two rules stems from high lift and confidence values and common LHS. Rule 16 is connected to Rules 1, 17, 23, 24, 30, and 36 and has a vibrant relationship network. According to Rule 16 and its connections, if {KOZAA} falls, {KOZAL}, if {VAKBN} falls, {HALKB}, if {EREGL} falls, {ISDMR} stock will fall. If {AKBNK} increases, {YKBNK}, if {YKBNK} increases, both {AKBNK} and {ISCTR} stocks will increase. This connection was formed due to high lift values. There are also other rules with the most connections: Rule 61, Rule 63, Rule 60, and Rule 65. Another issue is that Rules 31, 35, and 57 have no connection. Unconnected rules usually indicate that there is not a strong enough relationship in the data set or that they do not conflict with other rules.

3.5.2. Application 2

In Application 2, which includes the closing prices of BIST TUM-100 index stocks, 34 association rules were formed in 500 observation values of 340 stocks, 19 of which have 2 variables and 15 of which have 3 variables. According to the “is.redundant ()” function, 34 rules are “False”, meaning that these rules are not “redundant” and are meaningful on their own. According to the “is.significant ()” function, all rules have the value “True”, meaning that these rules represent a significant relationship in the non-randomly generated data set. According to the “is.maximal ()” function, 28 rules were “True”, meaning that these rules are more interesting and meaningful. 6 rules were “False”, meaning that these rules are covered by another rule.

Rule 1 {NATEN=0} ⇒ {ESEN=0}, which has the lowest Fisher’s exact p-value (3.35E-51), is statistically highly significant. Rule 1 also has the highest lift value with a value of 1.648. A detailed evaluation of Rule 1 according to functions and parameters is given in Table 6.

Table 6. Function and Parameter Evaluation According to Rule 1

Function, Parameter	Valid Values	Value	Explanation
Redundant	(T/F)	False	The rule is not redundant, it is meaningful on its own.
Significant	(T/F)	True	The rule is meaningful, it represents a relationship.
Maximal	(T/F)	True	The rule is among the maximal rules, it is interesting and meaningful.
Support	[0,1]	0.412	The rule occurs in 41.2% of all trades.
Confidence	[0,1]	0.837	ESEN is also present in 83.7% of the NATEN transactions.
Coverage	[0,1]	0.492	The proportion of transactions in which NATEN and ESEN move together constitutes 49.2% of all transactions.
Lift	[0,∞]	1.648	The probability of ESEN exhibiting a decrease or remaining stable is 164.8% higher than the probability of NATEN independently exhibiting a decrease or remaining stable.
Conviction	[0,∞]	3.026	It shows that when NATEN=0, the probability of ESEN=0 is low. That is, there is a strong relationship between these two events.
Leverage	[1,1]	0.162	The value of 16.2% indicates that this rule is not random and has a significant relationship in the data set.
Jaccard	[0,1]	0.701	It is the intersection and combination ratio that shows that the probability of the events NATEN=0 and ESEN=0 occurring together is quite high.
Phi	[1,1]	0.648	It shows that there is a positive and high correlation between NATEN=0 and ESEN=0 events.
Gini	[0,1]	0.210	A value of 21% indicates that the relationship in the rule is significant and that this relationship follows a certain order in the data set.
Cosine	[0,1]	0.824	It is the concordance rate that shows a high similarity between NATEN=0 and ESEN=0.
OddsRatio	[0,∞]	22.102	It is the concordance rate that shows that there is a very strong relationship between the events NATEN=0 and ESEN=0.
CohenKappa	[1,1]	0.648	It shows that there is a good agreement between the events NATEN=0 and ESEN=0.
Kulczynski	[0,1]	0.824	It shows that there is a strong relationship between the events NATEN=0 and ESEN=0.
Lambda	[0,1]	0.642	It is the value that measures the relational strength showing that there is a significant relationship between the events NATEN=0 and ESEN=0.
Fisher p.values	p<0.05	3.35E-51	The Fisher p-value indicates that the rule is highly statistically significant.

Source: Authors' Representation

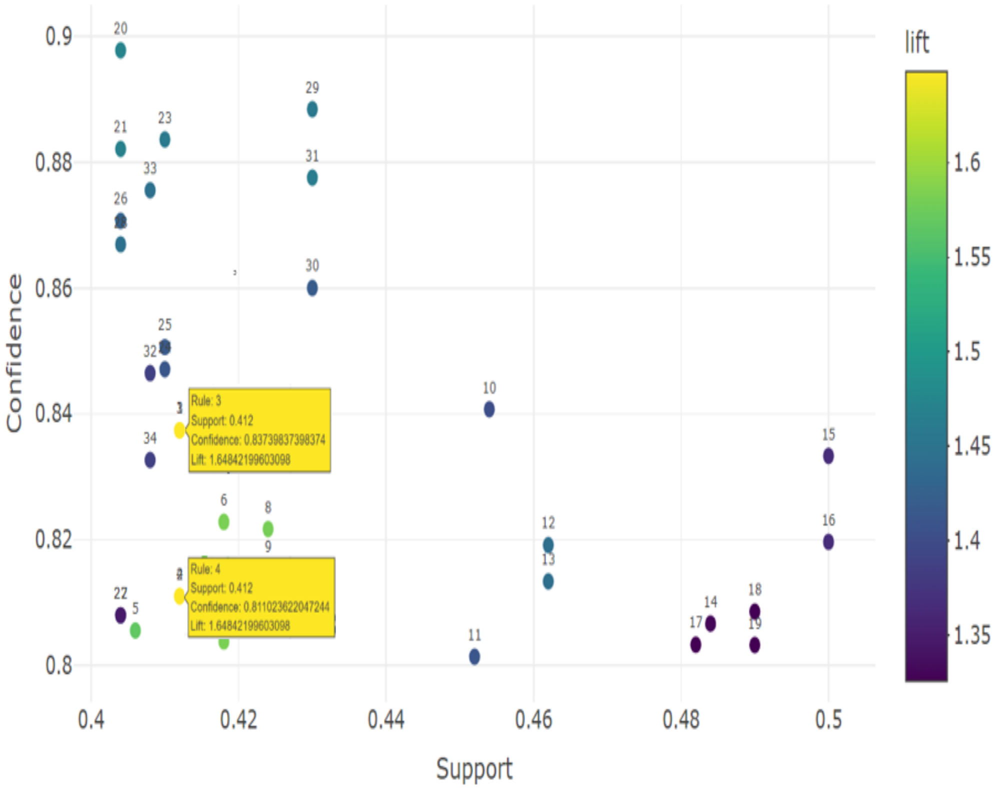
The scatter plot for Application 2 is given in Figure 7. The support and lift values are the same in the {NATEN} and {ESEN} associations that include rules 1, 2, 3, and 4 and are 0.421 and 1.648, respectively. In the rules of this association, "0" means decrease or remain the same, and "1" means increase, and these rules:

a. If {NATEN=0}, then {ESEN=0} with 83.7% (0.837) confidence.

b. If {ESEN=0}, then {NATEN=0} with 81.1% (0.811) confidence.
c. If {ESEN=1} shares increase, {NATEN=1} shares will also increase at 0.837 confidence level,
d. If {NATEN=1} shares increase, {ESEN=1} shares will also increase at 0.811 confidence level.

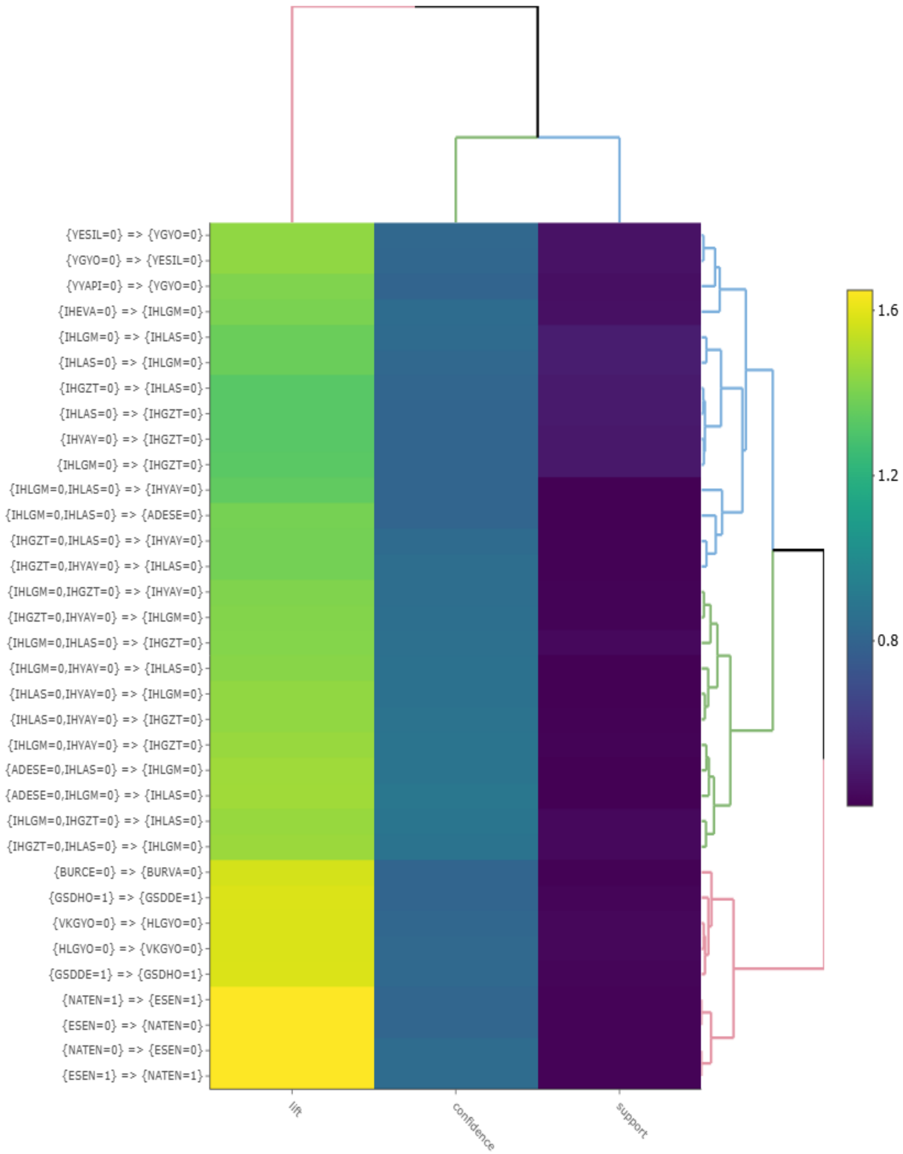
In the graph, these rules ((1-3) and (2-4)) are shown with yellow balloons, and the other rules are colored according to their lift values.

Figure 7. Scatter Plot (Application 2)



The grouped heatmap matrix graph created for the 34 association rules in Application 2 is given in Figure 8. 3 cluster groups were created according to the support, confidence, and lift values.

Figure 8. Grouped Heatmap Matrix Chart (Application 2)



Accordingly, Rules 1, 3, 4, and 2, which include {NATEN} and {ESEN} stocks shown in red in the dendrogram, have a lift value of 1.6 and above, and these stocks have both increasing and decreasing associations. In addition, {GSDDE} and {GSDHO}, {HLGYO} and {VKGYO}, {BURCE} and {BURVA} companies with lift values between 1.2 and 1.6

are also included in this cluster and consist of Rules 6, 8, 9, 7 and 5, respectively. For the 9 rules in this red cluster, {GSDDE} and {GSDHO} companies have increasing relationships; {BURCE} and {BURVA} and {HLGYO} and {VKGYO} companies have decreasing relationships. In the green cluster group, where 11 rules are included, there is a rare association of companies that can be predicted correctly. The 14 rules in the blue cluster have frequent associations in the data set.

The network diagram that illustrates which shares the association rules of Application 2 focus on is presented in Figures 9a and 9b, where Figure 9a shows a Circular Network Diagram of Association Rules and Figure 9b displays a Force-Directed Network Diagram of Association Rules. According to the diagram, the color of the Rule 1 node indicates that this rule has a high lift value, which indicates that the rule represents a significant and strong relationship. Rule 1 is linked to Rule 2. The fact that Rule 1 also highlights Rule 2 indicates that these rules should be analyzed together. Also, Rule 31, which has the most connections (7) in this graph, should be analyzed. The rules that have many connections like Rule 31 are: Rules 28, 29, 33. These rules and their connections should also be analyzed together. Also, Rule 9 and Rule 25 have no connections, meaning Rule 9 and Rule 25 do not have a strong enough relationship in the dataset.

Figure 9a. Circular Network Diagram (Application 2)

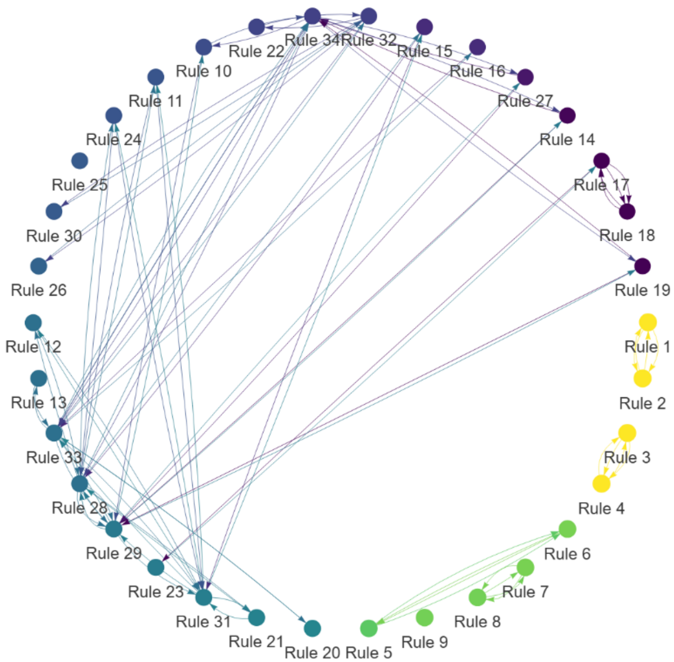
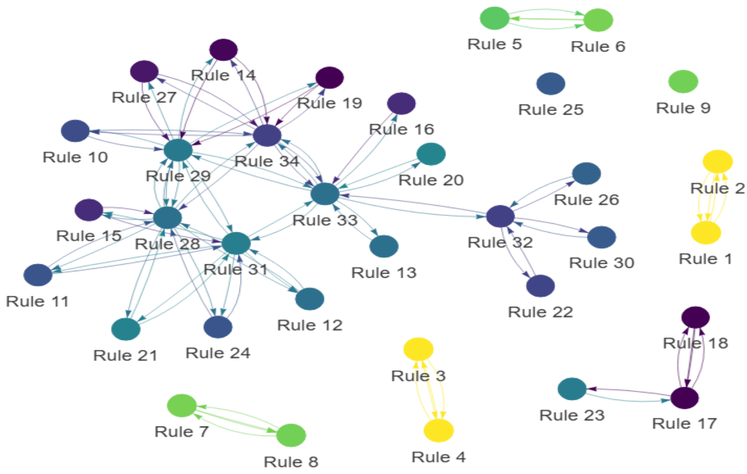


Figure 9b. Force-Directed Network Diagram (Application 2)



4. CONCLUSIONS

This study analyzed the association movements of stocks in BIST TUM (Application 1) and BIST TUM-100 (Application 2) indices with the Apriori algorithm and supported by data visualization techniques. The data sets for Application 1 and Application 2 consist of stock closing prices covering the period 10.01.2022/08.01.2024 obtained from the Finnet database. In the analysis performed in the R-Studio (4.3.3) programming language to obtain association rules, the basic parameters of the Apriori algorithm, support and confidence coefficients, were determined as 40% and 80% in this article, respectively. To make a more comprehensive evaluation using multiple criteria depending on these parameters, the following criteria were taken into account in the study: coverage, lift, leverage, conviction, jaccard, phi coefficient, odds ratio, gini, kappa, lambda, kulczynski, cosine. In the analysis where redundant (), maximal (), and significant () functions were also taken into account, 66 association rules were obtained in Application 1 and 34 in Application 2. All parameters in the analysis revealed meaningful and interesting rules. The importance of the rules can be ranked according to each of these parameters and the most important rule for each parameter may change. To prevent this, the importance ranking was done statistically with Fisher's exact test. Deciding which parameters are the best depends on the purpose and context of the analysis. In other words, if the rules are to be generalizable and widespread within the scope of the analysis, support is an important parameter, if they are reliable and accurate rules, confidence is an important parameter, and if they are strong relationships that exceed the independence assumption, lift is an important parameter. Also, using the traditional "support-confidence" pair for association mining in big data may not be ideal. What is desired in association rule analysis is to uncover unexpected, surprising associations. It can be quite difficult to find the most interesting association rules. For this, criteria that reveal interestingness can be used.

When the findings obtained within the scope of the research are evaluated from a financial point of view; it is thought to provide various information to stock investors in making investment decisions. After the applications, there are stocks whose movements were determined together. HALKB and VAKBN; AKBNK, ISCTR and YKBNK; KOZAA and KOZAL; AKBNK and ISCTR; EREGL and ISDMR; NATEN and ESEN; GSDDE and GSDHO; HLGYO and VKGYO; BURCE and BURVA are the main ones. When the shares with this co-movement are analyzed, the common points of HALKB and VAKBN are that they are public banks and both operate in the banking sector. While other public subsidiaries or affiliates are not included in the association rules, these two banks stand out. Again, AKBNK, ISCTR, and YKBNK, which are among the largest private sector banks in Turkey and lead the banking sector with their size, have been identified to act together. Their common feature is that they are banking sector companies. KOZAA and KOZAL have two strong ties. The first is that they operate in the mining sector. The second is that KOZAA owns 45 percent of KOZAL. In other words, KOZAL is a subsidiary of KOZAA. It can be considered that the capital structure and sector are the important reasons for the identified co-movement. A similar situation is observed in EREGL and ISDMR. Both are iron and steel companies and EREGL owns 95% of ISDMR. In Application 2, 3 more pairs of companies with similar shareholding structure and shareholding status were identified. The first one is NATEN and ESEN companies. Both are in the energy sector and NATEN owns 63% of ESEN. Second; GSDDE and GSDHO companies. Maritime transport is a common field of activity and GSDHO owns 68% of GSDDE. The last pair of companies are BURCE and BURVA. BURCE owns 40% of BURVA. Again, HLGYO, a subsidiary of HALKB (81%), and VKGYO, a subsidiary of VAKBN (70%), which are similar in terms of capital structure and field of activity, were found to act together. These examples do not mean that there will be co-movement in every subsidiary and partnership company. It should not be forgotten that the companies mentioned here were first identified to move together.

The findings obtained from the research can be used by both individual and institutional investors interested in portfolio management. The information presented here can be used by investors as follows: When the price of one of the companies whose co-movement is detected starts to rise, a long position can be considered as an investment opportunity if the other company has not yet risen. In case of a decline, a sell transaction can be made. In addition, investors who want to diversify their portfolios should not add both of the companies whose co-movement is detected to the portfolio. If one follows the price direction of the other, companies that move together should not be used for diversification to reduce portfolio risk.

This study, which could be considered an interdisciplinary research, aims to utilize data mining in financial decision-making processes. In this respect, the study may encourage researchers to engage in multidisciplinary work rather than conducting research within a single discipline. From this perspective, it is thought that it could serve as an example for academia.

In future studies, different criteria can be applied to a hybrid model based on the determination and clustering of price movements with a study that takes into account all BIST TUM and BIST TUM-100 stocks. Because the more data in the modeling step, the better results will be possible to obtain. The information obtained as a result of the study should not be considered as investment advice.

Araştırma ve Yayın Etiği Beyanı

Makalenin tüm süreçlerinde Manisa Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü Dergisi'nin araştırma ve yayın etiği ilkelerine uygun olarak hareket edilmiştir.

Yazarların Makaleye Katkı Oranları

Yazarlar çalışmaya eşit oranda katkı sağlamıştır.

Çıkar Beyanı

Yazarların herhangi bir kişi ya da kuruluş ile çıkar çatışması yoktur.

Acknowledgments

The second author of the article has utilized the knowledge gained from the “TÜBİTAK 2237-A BİDEB Statistical Modeling Techniques and Applications in Natural Sciences (Project No:1129B372300796)” and “TÜBİTAK 2237-A BİDEB Natural Science-Based Exploratory Data Analysis and Data Visualisation (Project No:1129B372400933)” projects, supported by TÜBİTAK, during the article process.

REFERENCES

Abidin, Z., Amartya, A.K., & Nurdin, A., 2022. Penerapan algoritma apriori pada penjualan suku cadang kendaraan roda dua (Studi Kasus: Toko Prima Motor Sidomulyo). *Journal Teknoinfo*, 16(2), 225-232.

Agrawal, R., & Srikant, R., 1994. Fast algorithms for mining association rules in large databases. *Proceedings of the 20th International Conference on Very Large Data Bases*, Santiago de Chile, 487-499.

Akpınar, H., 2017. *Data veri madenciliği ve veri analizi*, İstanbul.

Altunkaynak, B., 2022. *Veri madenciliği yöntemleri ve R uygulamaları*, Ankara.

Asur, F., Sevimli Deniz, S., & Yazıcı, K., 2020. Visual preferences assessment of landscape character types using data mining methods (apriori algorithm): The case of Altınsaç and Inkoy (Van/Turkey). *Journal of Agricultural Science and Technology*, 22(1), 247-260.

Ataman, G., & Kahraman, S., 2022. Comparing decision trees and association rules for stock market expectations in BIST100 and BIST30. *Scientific Annals of Economics and Business*, 69(3), 459-475. https://doi.org/10.47743/saeb-2022-0024.

Aydemir, E., 2019. *Weka ile yapay zeka makine öğrenmesi, Yapay sinir ağları-derin öğrenme*, Ankara.

Aydemir, E., & Yavuz, M., 2019. Mevsimlere göre ilaç satış verilerinin birliktelik analizi ile incelenmesi. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 3(1), 23-30. https://doi.org/10.33461/uybisbbd.582061.

Azis, N., Sucipto, P.A., Herwanto, A., Munthe, E.S., & Irwanto, D., 2024. Analysis of goods stock using the apriori algorithm to aid goods purchase decision making, *Sinkron: Jurnal Dan Penelitian Teknik Informatika,* 8(1), 533-541. https://doi.org/10.33395/sinkron.v9i1.13335.

Bayram, S.S., & Dündar, S., 2021. Türkiye'de banka şube lokasyonunun veri madenciliği ile analizi. *Uluslararası Bankacılık Ekonomi ve Yönetim Araştırmaları Dergisi*, 4(1), 34-52. https://doi.org/10.52736/ubeyad.936519.

Borsaistanbul.com. https://borsaistanbul.com/tr/.

Cemaloğlu, N., & Duykuluoğlu, A., 2020. *Sosyal bilimlerde veri madenciliği*, Ankara.

Çelik, C., Aladağ, Z., & Alkan, A., 2023. Veri madenciliği ve stok tutma süresi teknikleri ile depolar için dinamik ürün analizi modelinin geliştirilmesi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 29(2), 172-182. https://doi.org/10.5505/pajes.2022.29974.

El Mahjouby, M., Taj Bennani, M., Lamrini, M., & El Far, M. (2024). *Association rules forecasting for the foreign exchange market. International Journal of Electrical and Computer Engineering (IJECE), 14*(3), 3443-3454. https://doi.org/10.11591/ijece.v14i3.pp3443-3454.

Erden, C., 2021. *Python ile veri madenciliği*, İstanbul.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P., 1996. From data mining to knowledge discovery in databases, *AI Magazine*, 17(3), 37-54. https://doi.org/10.1609/aimag.v17i3.1230.

Hahsler, M., & Hornik K., 2007. New probabilistic interest measures for association rules. *Intelligent Data Analysis*, 11(5), 437-455. https://doi.org/10.3233/IDA-2007-11502.

Hahsler, M., 2015. A probabilistic comparison of commonly used interest measures for association rules. United States. Southern Methodist University. https://mhahsler.github.io/arules/docs/measures.

Hahsler, M., Gruen, B., & Hornik, K., 2023. *Mining association rules and frequent itemsets,* 'arules'. https://doi.org/doi:10.18637/jss.v014.i15.

Han, J.W., Kamber, M., & Pei, J., 2012. *Data mining concepts and techniques*. 3rd Edition, Morgan Kaufmann Publishers, Waltham.

Hand, D.J., Mannila, H., & Smyth, P., 2001. *Principles of data mining*, The MIT Press, Cambridge.

İnce, A.R., & Alan, M.A., 2022. Yatırım alanlarının planlamasında veri madenciliğinden yararlanılması üzerine bir çalışma, *Ekonomi, Politika & Finans Araştırmaları Dergisi*, 7(1), 1-15. https://doi.org/10.30784/epfad.1003459.

Karaatlı, M., Kocabıyık, T., Yalçıner, Çal, D., & Çolak, M., 2021, BIST-30 endeksinde yer alan payların ortak hareketlerinin veri madenciliği kapsamında birliktelik kuralı ile incelenmesi. *Finansal Araştırmalar ve Çalışmalar Dergisi*, 13(25), 548-571. https://doi.org/10.14784/marufacd.976609.

(Submission: 29.08.2025 / Published:29.12.2025)
Doi Number: 10.18026/cbayarsos.1773973
Type of Research: Research Article

Karasu, B., & Doğan, O., 2020. Birliktelik kuralları madenciliği ve ilginçlik ölçümleri: bir vaka çalışması. *JOBDA*. 3(2), 94-107. https://doi.org/10.46238/jobda.811464.

Kartal, B., Sert, M.F., & Kutlu, M., 2022. Determination of the world stock indices co-movements by association rule mining, *Journal of Economics, Finance and Administrative Science,* 27(54), 231-246. https://doi.org/10.1108/JEFAS-04-2020-0150.

Kocabıyık, T., Dağ, O., & Karaatlı, M., 2021. Borsa istanbul endekslerinin birlikte hareketi: fp growth algoritması ile bir uygulama. *Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi (IJBEMP)*, 5(2), 659-672. https://doi.org/10.29228/ijbemp.52518.

Kocabıyık, T., Karaatlı, M., Özsoy, M. & Özer, M.F., 2024. Cryptocurrency portfolio management: a clustering-based association approach. *Ekonomika*, 103(1), 25–43. https://doi.org/10.15388/Ekon.2024.103.1.2.

Liu, W., 2024. Research on shareholder association analysis based on apriori algorithm of machine learning. *In 2024 IEEE 7th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Chongqing, China, 7, 1810-1813. https://doi:10.1109/IAEAC59436.2024.10503928.

Oğur, G., & Borat, O., 2023. Müşteri profili ve alışveriş hareketlerini belirlemede rfm analizi ve birliktelik kuralları analizi: perakende sektöründe bir uygulama. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 22(44), 359-376. https://doi.org/10.55071/ticaretfbd.1327068.

Oğuzlar, A., 2004. Veri madenciliğinde birliktelik kuralları. *Öneri Dergisi*, 6(22), 315-321. https://doi.org/10.14783/maruoneri.678958.

Özkan, Y., 2016. *Veri madenciliği yöntemleri bilgisayar bilimleri ve bilgisayar mühendisliği*, İstanbul.

Pekdemir, R., Süer, A.Z., Ercan, M., & Önder, E., 2023. Analyzing and understanding social marketing activities of turkish banking corporations: comparisons and association rule learning, *Journal of Accounting and Finance*, 23(3), 102-114. https://doi.org/10.33423/jaf.v23i3.6354.

Teker, T., & Konuşkan, A. (2022). Fan Token Fiyatlarında Birliktelik Etkisi. *Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi,* 6(2), 359-376. . https://doi.org/10.29228/ijbemp.65914.

Uslu, M., 2018. Birliktelik kuralları analizi (association rules analysis). https://www.veribilimiokulu.com/associationrulesanalysis/.

Ünsal, Ö., 2020. Veri madenciliği teknikleri ile hisse senetleri arasındaki fiyat etkileşimlerinin belirlenmesi. *Mühendislik Bilimleri ve Tasarım Dergisi*, 8(5), 106-112. https://doi.org/10.21923/jesd.834105.

Xie, H. (2021). *Research and Case Analysis of Apriori Algorithm Based on Mining Frequent Item-Sets*. *Open Journal of Social Sciences, 9*, 458-468. https://doi.org/10.4236/jss.2021.94034.