Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/180206

Using Data Analytics for Evaluating Social and Economic Impacts of Post-COVID Pandemic

Aline Abboud[1], *Hassan Harb*[2,*], *Alaaeddine Ramadan*[3], *Chamseddine Zaki*[2], *Louai Saker*[2], *Moustafa Ibrahim*[2] *and Abbass Nasser*[4]

[1]Faculty of Sciences, Lebanese University, Hadath, Lebanon
[2]College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait
[3]Department of Computer Engineering, College of Engineering, Ahlia University, Manama, Bahrain
[4]Faculty of Arts and Science, American University of Culture and Education (AUCE), Beirut, Lebanon

Received: 7 Oct. 2023, Revised: 21 Dec. 2023, Accepted: 23 Jan. 2024
Published online: 1 Mar. 2024

Abstract: The global spread of COVID-19 exposed flaws in healthcare systems worldwide, hindering effective response to the outbreak. Beyond a health crisis, it disrupted societies, causing widespread business closures and economic losses. Urgent socio-economic solutions are crucial to prevent prolonged suffering and safeguard lives and livelihoods. In this paper, we propose a technique that follows typical algorithms to understand the entire theory behind the effect of COVID-19 on socioeconomic levels. By convention, the technique we are mentioning consists of six stages: Data Collection, Data Cleansing, Data Transformation, Data Discretization, Correlation Analysis, and Association Rules. We collected data from approximately 500 participants using a survey addressing questions about their socioeconomic status in light of the COVID-19 pandemic. All stages are related, starting from the Data Collection Stage which is distributing the survey and retrieving data using Google Forms. Next, follows the Data Cleansing stage which fills any gaps in the raw data coming from the CSV File and throws this newly cleansed data to the next stage, the Data Transformation, which does the necessary transformations to apply the algorithms. Also, the Data Discretization stage aims to discretize non-meaningful data into intervals of meaningful data. These four stages allow the data to be ready for the Correlation Analysis stage which shows the correlation between all variables in the questionnaire. Finally, leveraging the Apriori algorithm—an optimized version of Association Rules—we extract significant if-then patterns from the processed data. All procedures were executed using the R Programming Language and its associated libraries.

Keywords: COVID-19, Socio-Economic Impact, Correlation Analysis, Association Rules, Apriori Algorithm

1 Introduction

The healthcare sector plays a pivotal role in supporting other sectors, hence the necessity of strengthening this sector by matching effective preventive interventions with appropriate treatments. Therefore, the government should mobilize and allocate finances and investments in the health system. Many countries have reserved significant budgets and raised remarkable funds in support of healthcare functionality and performance. Those funds were allocated for manufacturing medical equipment, vaccines, and drugs as well as for the provision of medical insurance. Furthermore, a significant amount of funds was granted to scientific research.it is worth mentioning the prominent role of the World Health Organization (WHO) in providing support to countries in ways that better respond to their needs and set the necessary health measures in the face of any emerging disease. Nevertheless, the appearance of the COVID-19 pandemic has revealed a deficit in the funding and preparedness of the healthcare system whether nationally or internationally, even in countries with the most sophisticated healthcare system. The COVID-19 pandemic has expanded health economics literature to include government policy and health system innovations. Recent research [1] examined economic evaluations of government policies to reduce COVID-19 transmission, broad health system innovations, and models of care. This can aid pandemic policymaking and future economic evaluations. On the 30th of January 2020 WHO declared

* Corresponding author e-mail: hassan.harb@aum.edu.kw

the outbreak of COVID-19 as a global pandemic and since then, a daily epidemiological update illustrated the fast spread of this virus globally; On the 2nd of May 2023, the number of confirmed cases has risen to around 728.830 million with 6.866 million deaths [2]. Even though the majority of people affected by the virus have mild symptoms and a significant number have experienced severe illness and even death. Beyond the tragic health threats of the COVID-19 pandemic, its impacts are far-reaching and long-term. The COVID-19 outbreak has globally threatened all societies on all levels: public health, economy, education, and environment. It has burdened societies and economies with heavy adverse outcomes. Millions of businesses have faced existential threats with billions of workers at risk of losing their livelihoods. The United Nations Trade and Development Agency (UNCTAD) estimates that global economic growth is projected to decelerate to 2.5% in 2022, followed by a further decline to 2.2% in 2023. The global deceleration in economic activity would result in a real gross domestic product (GDP) that remains lower than its pre-pandemic trajectory, resulting in a loss of over $17 trillion, approximately equivalent to 20% of the world's total income [3]. Consequently, there has been a decline in global economic growth, reaching approximately 2.6% by the conclusion of 2022, and is projected to further decrease to 2.0% in 2023. This anticipated rate represents the lowest level of growth observed since the year 1970 [4]. The COVID-19 outbreak touched all segments of the population differently. It has shed light on the underlying disparities among the most vulnerable classes of society: people living in poverty circumstances, the elderly, and people with disabilities. It is expected that poverty, inequalities, and social discontent will grow on a global scale. In this context, evaluating the effects of the COVID-19 crisis on societies, economies, and vulnerable groups is pivotal for tailoring government responses to recover from the crisis. In this paper, we proposed a technique aiming to determine the socioeconomic implications of the COVID-19 pandemic among the people. By convention, the technique that we are mentioning consists of six stages starting with the first stage which is collecting the data that is related to the population, then second, third, and fourth stages by pre-processing them respectively by cleansing, transforming, and discretization the data then on the fifth one we apply the analysis process with the Correlation between attributes. Finally, the sixth stage is applying the Apriori algorithm to extract interesting if-then facts from the data we have in hand. We assessed the efficiency of this technique built on real collected data from more than 500 individuals while the results show a clear understating of the effects of this pandemic on the socioeconomic core. The remainder of this paper is organized as follows. An overview of the methods suggested for researching the effects of COVID-19 is provided in Section 2. We outline each data stage and present the framework in section 3. The simulation and the obtained results are presented in Section 4. Section 5 provides a conclusion to the paper as well as guidelines for further research.

2 Related Work

Researchers are overloaded with COVID-19 papers; an analysis of academic publications devoted to the virus reveals a sharp increase in the number of papers about COVID-19 surged in 2020; in a mere half-year, prominent repositories were overrun with study articles, correspondence, reviews, notes, and editorials. Of all COVID-19-related articles on Scopus and Web of Science, research papers made up 48% and 37% of the total, respectively [5]. Researchers are mobilizing to study the impacts of this pandemic on different aspects such as Psychologists, healthcare is shifting focus and obtaining funding to study the mental health and behavioral science aspects of the pandemic [6,7,8], economical [9,10,11], agriculture [12], tourism [13], environment [14], educational [15,16,17], etc. The government has a major role to play in the socio-economic transition of society as a result of the coronavirus pandemic. Many researchers were performed in light of these incidents. Unfortunately, this outbreak caused a massive economic shock across the world and most of the research techniques related to the Covid-19 pandemic have shown how this outbreak has alarming social and economic implications for individuals as well as the population. It has also resulted in major demographic changes, unemployment, and the closure of economic activities. Many vulnerable communities around the world have experienced serious social and livelihood effects as a result of the spread of the coronavirus. Millions of companies have faced extinction, and billions of employees are at risk of losing their jobs. The study of [18] aims to analyze these socio-economic effects to decide how the widespread is causing different issues to the poverty-stricken. A microeconomic model is developed to determine the direct impact of dividing the family's income, savings, expenditure, and scarcity as a first step in assessing the household-level impacts of COVID-19 at a territorial scale. Using the San Francisco Bay Area as a case study, the study's main conclusions show that COVID-19 will have a substantial negative impact on the economy in the absence of adequate social protection, moreover, government aid will lead to a decrease in the volume and duration of the crisis. However, the economic effect is spatially diverse, with some areas being hit harder than others and taking more than a year to recover. In [19], the study conducted by the authors delves into the economic impact of COVID-19 on small businesses, revealing financial insecurity in over 5,800 of them. The findings reveal that if the recession continues longer than four months, companies in especially vulnerable sectors, such as restaurants, tourism, and personal services, may find it extremely difficult to remain in business. These results stress the value of well-designed and long-term economic and public health strategies that can securely shorten the economic shutdown. In [20], 748 small and medium enterprises (SMEs) in the furniture industry in Malaysia participated in an online survey. The main aim was to look into the effect of the pandemic and subsequent Movement Control order (MCO) on various sides of SMEs' business, as well as make recommendations to the government on how to help them. It was proposed that the government take steps to assist SMEs in controlling their cash flow and relaxing constraints in order to promote the critical supply-chain start-up. According to the survey, most SMEs understand the significance of automation and technology adoption. Respondents agree that the transition to Industry 4.0 is imminent as a way to improve their strength in the face of potential uncertainties. The authors of [21] examine the role of government. A semi-structured questionnaire and a non-probability snowball sampling methodology were used to perform an online survey; the researcher obtained responses from 100 individuals. The study explores various aspects of people's socioeconomic status, as well as obstacles to improving their wages, GDP, consumption, and investment levels, as well as whether the government is making a substantial effort to improve people's living standards following the pandemic. A limited effort has been made to comprehend the people's tension and anxiety during the pandemic, as well as how the government can assist in the transition of society. In [22], the authors aim to determine the socioeconomic consequences of the COVID-19 pandemic, especially on various organizations and individual consumption patterns. On the social front, researchers will investigate how the pandemic's social distancing, panic purchasing, and other prevention measures have changed people's attitudes in their relationships with one another. As a result, interviews with family members, laborers, and employers were conducted, from a socioeconomic perspective, this study looked at how global pandemics influence human actions. The paper [23] examined the COVID-19 outbreak patterns in 14 Latin American countries. The study shows significant discrepancies between reported COVID-19 cases and the epidemiological situation, highlighting the need for accurate data handling and continuous surveillance in epidemic management. The results also show that country size does not affect COVID-19 cases or deaths, suggesting other factors are involved. Quarantine has reduced real-time reproduction, but infection rates have increased after daily activities resume. These findings demonstrate the difficulty of balancing public health with economic and social activities. This research informs pandemic control strategies and decision-making.

3 Proposed Mechanism

The pandemic of COVID-19, which began more than three years and seven months ago, has had a lasting impact on nations around the globe. Although significant progress has been made in overcoming the pandemic and many nations are gradually returning to normalcy, the pandemic's effects continue to linger. During the height of the virus's spread, the international community faced an unprecedented emergency. Despite the introduction of vaccines more than a year ago, the socioeconomic burden of COVID-19 persists. It is essential to assess the long-term socioeconomic effects of the pandemic, including the fears and consequences faced by individuals and society, as well as the increased levels of poverty. This study aims to provide a comprehensive understanding of how the outbreak has affected Lebanese citizens, in order to formulate effective strategies and policies to confront any future pandemic. As depicted in Fig. 1, the research consists of six stages.

Fig. 1: study stages

3.1 Data Collection Stage

The first step of our work is to collect the data from the questionnaire. For that purpose, a survey is performed with Google Forms launched between November and December 2022, consisting of 20 questions about different sections of the social and economic impact of the COVID-19 pandemic. Since this is a sensitive topic, the questions were filled by anonymous users to guarantee the user's confidentiality and privacy. More than 500 people participated in this study, which was directed towards the Lebanese population in various regions. The survey's questions were carefully crafted and divided into three primary categories:

- *General questions:* consist of 8 questions related to the participant's gender, psychological education, age, social status, hobbies, personal monthly salary, family members count, and family monthly salary.
- *COVID-19 Information Questions:* consists of 3 questions about the method of receiving COVID-19 updates, quarantine dedication, and prevention level.
- *Psychological-based questions* during COVID-19 consist of 9 questions covering the level of fear of the virus, the content of this fear, the psychological health care level, the level of physical health fear, relatives'

physical health fear, and monetary situation fear because of the virus and finally the level of fear from viruses in general.

From a detailed view, questions are provided as single and multiple-choice. Some of these single-choice questions are data to be transformed or discretized such as the level of fear. Now, of course, questions like salary are also a single choice but with a provided range of value, and so on. Although data seems kind of simple, it is a bit complex to work with and identify the results but provides a better understanding of the participants and their actual state as well as contributes to a deep knowledge of the social and economic impact of the COVID-19 pandemic.

3.2 Data Cleansing Stage

Given that the entire survey relies on single and multiple-choice questions, most of which are mandatory, data cleaning is facilitated by the stringent control inherent in the questionnaire. Addressing missing values involves filling in approximately 60 answers for the remaining non-required questions, rather than excluding them from the dataset. Instead of removing empty values, we opt for a strategy of replacing them with default values specific to each question. For instance, if the non-required ”Fear Content” question is left unanswered, it is replaced with ”Nothing,” signifying an absence of fear. This approach ensures that the data row is retained, preventing loss due to missing answers. Throughout the data cleansing process, we carefully assess the potential impact of using default values on association results. Sensitivity analyses are conducted, comparing association outcomes with different default value choices. This scrutiny enables us to conserve the robustness of the associations against variations in default values. Consequently, the use of default values enhances the analysis and does not introduce unexpected patterns.

3.3 Data Transformation Stage

In the survey data, responses are initially recorded in textual formats. However, to facilitate meaningful calculations and analyses, a crucial step involves transforming the textual representations into numerical values, particularly when dealing with categorical variables. For example, consider personal and family monthly salaries presented in a range format, such as ”Between $1,000 and 2,000$.” To enable numerical calculations, these ranges are conventionally transformed into the numerical average 1500. Similarly, numerical values in character format, like ”Family Members Count”, transform into an integer type for compatibility with analytical processes. Additionally, static single-choice questions, like ”Gender,” are transformed into factors, assigning numerical values (0 and 1) to the respective levels (”Male” or ”Female”). This transformation ensures the data is in a format conducive to robust statistical analyses and algorithmic applications.

3.4 Data Discretization Stage

In this phase, we generate essential data for subsequent analyses by discretizing numerical information into meaningful categories. Indeed, the data discretization process results in the addition of a few new columns that contribute to a richer dataset. One significant addition is the computation of the average family salary derived from the responses to ”Family Monthly Salary” and ”Family Members Count.” This calculation, outlined in equation 1, results in the creation of a new column that encapsulates a more nuanced understanding of economic dynamics within the dataset.

$$F(f,n) = \frac{f}{n} \tag{1}$$

In Equation 1, where ”f” represents the ”Family Monthly Salary” and ”n” signifies the ”Family Members Count,” the Average Family Salary is computed and subsequently discretized into five distinct intervals, each denoting a specific socioeconomic category:

– Very Poor (numerically represented as ”0”)
– Poor (numerically represented as ”1”)
– Regular (numerically represented as ”2”)
– Rich (numerically represented as ”3”)
– Very Rich (numerically represented as ”4”)

Following a similar discretization strategy, the process is extended to calculate the “Level of Fear regarding physical health”, “Level of Fear on relatives’ physical health”, “Level of Fear due to lockdown”, “Level of Fear on the monetary situation” and “Level of Fear from viruses in general”. For instance, the measure “Level of Fear regarding physical health” is discretized into three distinct categories:

– I have no fear regarding my physical health.
– I have a regular level of fear concerning my physical health.
– I have a severe fear regarding my physical health.

The preceding four stages, focused on data gathering and meticulous preparation before the application of analysis techniques, can be succinctly summarized in the visual representation provided in Fig. 2.

3.5 Correlation Analysis Stage

We applied rigorous statistical methods to examine the intricate relationships among various variables that could potentially impact the perceived level of fear. Specifically, the Pearson and Spearman correlation coefficients served

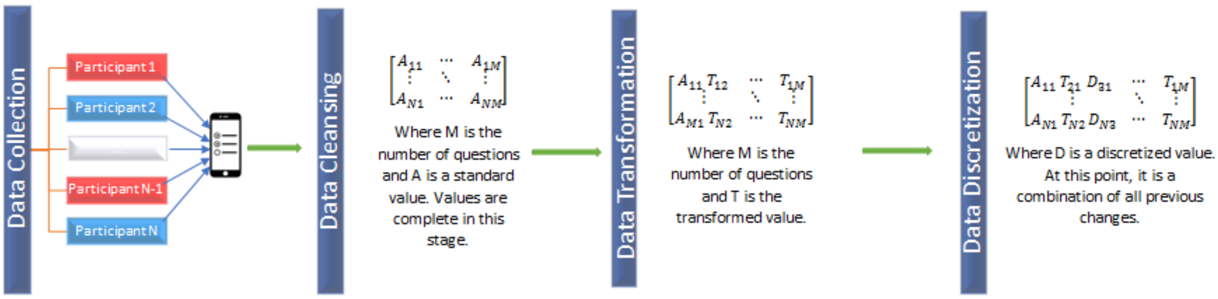

Fig. 2: Study stages

as robust tools for assessing these relationships. Additionally, we leveraged multiple linear regression techniques, employing forward selection methods to refine the precision of our analyses. The correlation coefficient should exhibit a discernible deviation from zero, signaling a meaningful positive or negative relationship, and ideally approaching values of 1 or -1.

–Pearson Correlation:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \tag{2}$$

r represents the correlation coefficient; x_i represents the values of the x-variable in a sample; $\bar{x}$ represents the mean of the values of the x–variable. y_i = values of the y -variable in a sample, and $\bar{y}$ represents the mean of the values of the y-variable

–Spearman Correlation:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3}$$

ρ = Spearman's rank correlation coefficient
d_i = difference between the two ranks of each observation
n = number of observations

3.6 Association Rules Stage

Association rules are "if-then" statements that help to show the probability of relationships between data items, within large datasets. Using this method allows us to discover various interesting relationships between variables, which is the target of the survey. Discovering the relationship between poverty level and COVID-19 on psychological levels is the key here. The advanced and optimized version of the association rules algorithm is the Apriori algorithm (Algorithm 1), which goes as follows:

Algorithm 1 Apriori

Require: D: transaction database, Min_sup: the minimum support threshold
Ensure: frequent itemsets

```
L_1 = find_frequent_1-itemsets(DB)
for (k = 2; L_{k-1} = φ; k++) do
    C_k = Apriori_gen(L_{k-1})
    for each transaction t ∈ DB    //scan DB for counts do
        C_t = subset(C_k, t)    //get the subsets of t that are candidates
        for each candidate c ∈ C_t do
            c.count ++
        end for
    end for
    L_k = {c ∈ C_k suchthat c.count ≥ min_sup}
end for
return L = ⋃_k L_k
Procedure Apriori gen(L_{k-1}: frequent(k − 1)-itemsets)
```

Parameter	Description	Values
τ	Correlation Thresholds	0.3, 0.4, 0.5
σ	Association Rules Support	0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
ϕ	Association Rules Confidence	0.6, 0.9

Table 1: Simulation configuration.

4 Performance Evaluation

In this section, we show the performance of our mechanism in order to understand the impact of the COVID-19 pandemic on psychology and financial situation. As introduced earlier, our program is implemented in the R programming language, and we used the data collected from a survey submitted by more than 500 persons as described in section 3.1. We run this system on a Windows 10 operating system that possesses an Intel Core i5 CPU and 4 GB of RAM. Table 1 provides the parameters used in our simulations.

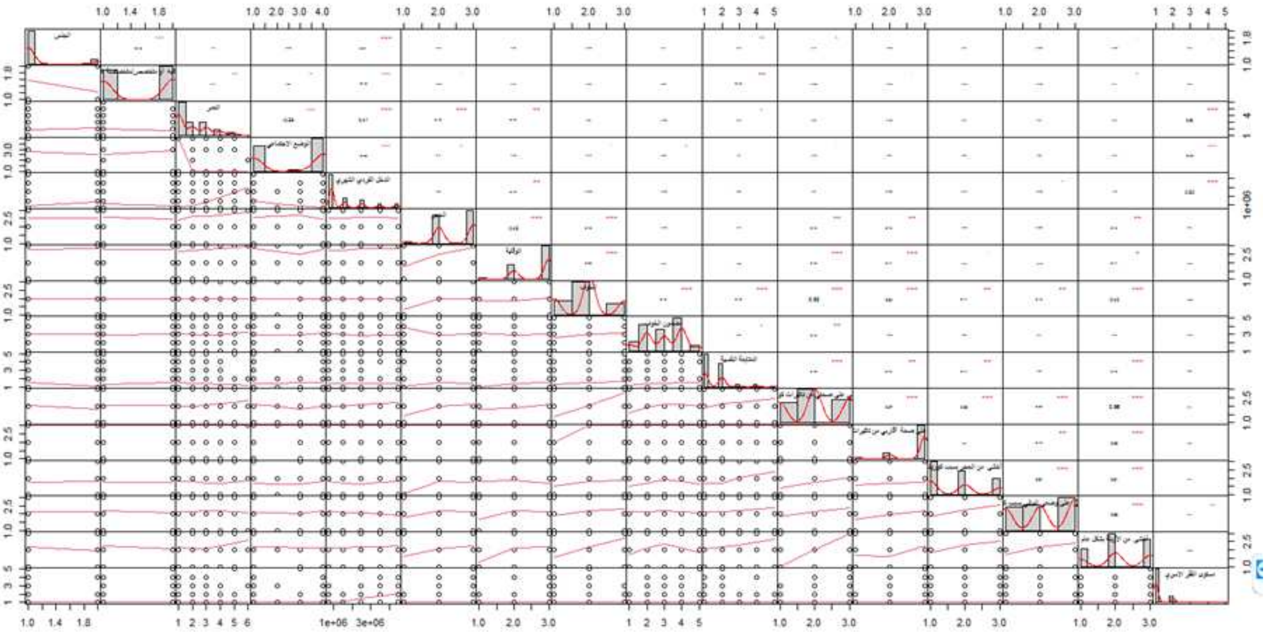

Fig. 3: Global correlation Chart

4.1 General Study of Correlation

In this section, we perform a correlation analysis on pretty much most of the variables with each other, in a global-like view to better understand which factor affects the other, and so on. As we know, to have a real correlation, the Correlation Value should be different than zero, either positively or negatively, with a value as close as possible to 1 or -1. In other words, we can detect a correlation using scatter charts rather than correlation values analysis. If the line of the chart is strictly positive (x increases when y increases), or if this line is strictly negative (x decreases when y decreases), then we can assume that a correlation is present. Otherwise, if this line is slightly or completely horizontal, we may say that there is just a little bit of correlation which is not enough or no correlation at all. The following performance correlation combination chart of all factors provides a better way to understand the correlation charts in a global view (Fig. 3). The diagonals are histograms indicating if this factor has normally distributable data. If a Bell-Curve is detected, it means that this data is normally distributed, else it is not. For instance, factors 8, 11, 14, and 15 have their data normally distributed as a result of a Bell-Curve. The lower triangular part shows the scatter charts of the correlation between factor i and factor j, while the upper triangular part shows the correlation values between them. These values are not clear here due to the size of the chart, but we will see these values in a bit. For example, for scatter at $i = 15$, $j = 11$, there is a positive correlation. Similarly, for $i = 2$, $j = 1$ there is a negative correlation. Finally, for $i = 13$, $j = 2$ there is no correlation at all.

4.2 P-Values

Every correlation analysis involving two factors or variables, two critical values come to the forefront: the correlation coefficient and the P-Value. The P-Value (Fig. 4) is a numerical representation that signifies the validity of the observed correlation against a predefined threshold. This concept is grounded in statistical studies that yield outputs to ascertain the veracity of the correlation value. The analytical process is underpinned by two primary hypotheses:

– *Null Hypothesis:* Positing that there is no perceptible difference between the correlation coefficient and 0. In the population, there is no significant linear relationship (correlation) between studied variables.
– *Alternative Hypothesis:* Contending that there exists a noteworthy difference between 0 and the population correlation coefficient. In the population, a significant linear relationship prevails between studied variables.

When the P-Value falls below 5 percent (0.05), we ”reject” the Null Hypothesis, signifying the presence of a correlation. Conversely, if the P-Value exceeds this threshold, the Null Hypothesis is accepted, indicating a lack of correlation. The symmetrical heat matrix presented below illustrates the P-values for all considered factors. The intensity of the color reflects the likelihood of correlation; darker shades suggest less likelihood, while lighter shades indicate a stronger possibility of correlation. Values below 0.05 signify a correlation. For example, the zero values above unequivocally assert that the correlation value is significantly different from 0.

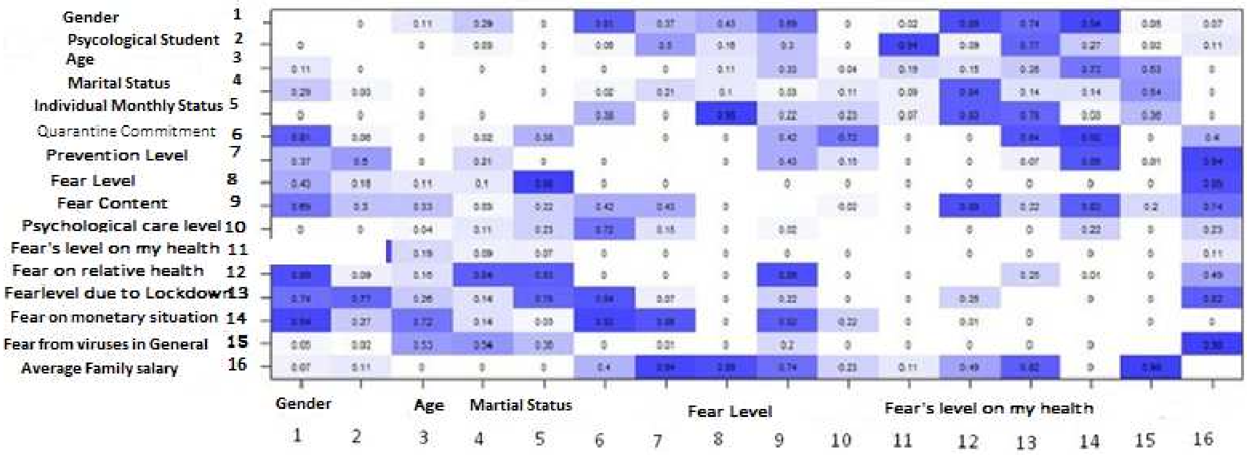

Fig. 4: P-Values

Conversely, higher values confirm that the correlation value is either 0 or near it.

4.3 Thresholding Correlation Values

The correlation coefficient is a statistical measure indicating the strength of the relationship between the relative movements of two variables, with values ranging between -1.0 and 1.0. A correlation of -1.0 shows a perfect negative correlation, while a correlation of 1.0 shows a perfect positive correlation. In this section. In order to identify the most highly correlated factors throughout our correlation analysis, we employ a thresholding mechanism on our data, illustrated in Fig. 5. This approach enables us to pinpoint and concentrate on variables that display significant and noteworthy correlations, streamlining our analysis for a more targeted examination of key relationships.

–No Threshold: the following symmetric correlation heat matrix shows all the correlation values. The stronger the red color is, the better the negative correlation is. Similarly, the stronger the blue color is, the better the positive correlation is. Whitish values have their factors considered as non-correlated (Fig. 5(a)).

–Threshold of τ = 0.3: the following symmetric correlation heat matrix shows the correlation values that are at least 0.3 (Fig. 5(b)).

–Threshold of τ = 0.4: the following symmetric correlation heat matrix shows the correlation values that are at least 0.4 (Fig. 5(c)).

In dissecting the findings derived from the survey encompassing around 500 participants, a discernible trend emerges: the fear of COVID-19, in and of itself, does not exert a direct impact on individuals' financial situations. Instead, what becomes apparent is that the apprehension associated with staying at home during quarantine due to the pandemic is the driving force behind financial concerns. This observation aligns with the earlier-discussed theory – individuals experiencing fear of financial implications are those who, when confined to home without employment, are more likely to encounter financial challenges. In essence, it is the fear associated with the quarantine lifestyle that appears to catalyze apprehensions about one's financial well-being.

4.4 Association Rules Analysis

The Apriori algorithm, known for generating a substantial number of association rules, necessitates careful control of its two primary variables: support and confidence. To meticulously explore the impact of different combinations of these parameters, we conducted a comprehensive analysis, varying the values of support (σ) and confidence (ϕ) within defined ranges.

–Confidence ϕ =0.6, variable support σ: we set a confidence threshold (ϕ) of 0.6 while systematically varying the support variable (σ) within the range [0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]. Fig. 6(a), visually encapsulates the impact of these selected values on the generation of association rules.

–Confidence ϕ =0.9, variable support σ: we set a confidence threshold (ϕ) of 0.9 while systematically varying the support variable (σ) within the range [0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]. Fig. 6(b), visually encapsulates the impact of these selected values on the generation of association rules. We demonstrate that lower levels of anxiety manifestations are

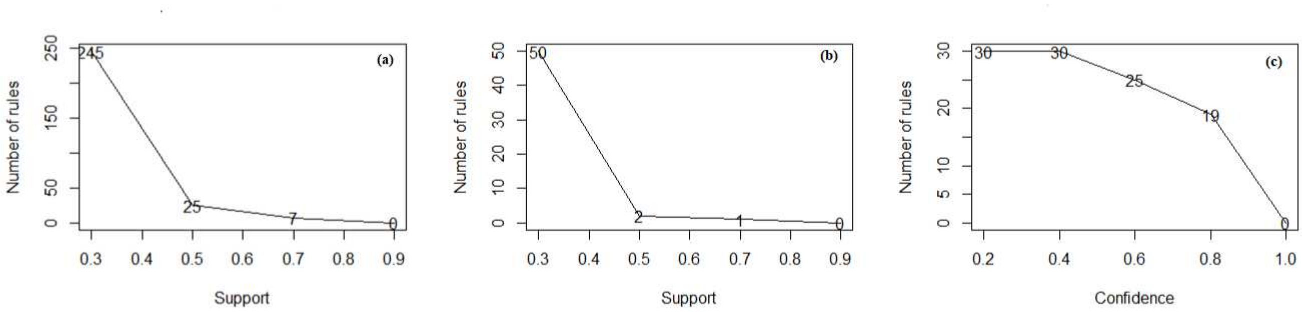

Fig. 5: (a) All Correlations with No threshold. (b) Correlation with threshold =0.3. (c) Correlation with threshold =0.4.

correlated with activities such as working, playing, walking, reading, socializing, and relaxing in the third set of clusters. These people are women with a specialty in psychology who engage in these consoling activities. It demonstrates that these individuals place a strong emphasis on the psychological benefits of these activities for anxiety management.

– Support σ = 0.5, Variable Confidence ϕ: we set support at 0.5. while systematically varying confidence across the range [0.2, 0.4, 0.6, 0.8, 1.0]. Fig. 6(c).

We proceed by selecting specific instances of these association rules to scrutinize the "if-then" statements, aiming to conclude the relationships between survey variables. A sneak peek into the examples, for Confidence ϕ =0.6, variable support σ =0.3 as shown in Fig. 7 we get the 245 association rules, a part of which is illustrated in Fig. 7. For example, in the realm of association rules, a significant discovery highlights the inherent link between a disadvantaged economic situation and the widespread apprehension regarding viral threats. What makes this correlation particularly compelling is the absence of an opposite rule, further reinforcing the credibility of the initial findings. Upon deeper investigation, overarching rules begin to emerge, shedding light on the direct interaction between various types of fear, especially those induced by global pandemics like COVID-19, and their profound impact on individuals. This impact encompasses both psychological well-being and pertinent physical health concerns.

5 Conclusion

The world has completely collapsed due to COVID-19. Our way of living, relating to one another, working, and communicating have all been impacted. Unfortunately, COVID-19 is leaving the already vulnerable people at a further disadvantage. It will affect societies all around the world for years to come; therefore, governments must have clear and available information in order to make a good decision to deal with this pandemic. In this paper, we use a six-stage technique to summarize the socioeconomic effects of COVID-19 on different facets of the global economy: Data Collection, Data Cleansing, Data Transformation, Data Discretization, Correlation analysis, and Association Rules. Mainly our data was collected from more than 500 persons followed by 4 processes to allow the data to be ready for the Correlation Analysis and then to apply the Apriori algorithm to extract interesting if-then facts from the data we have in hand. We could enhance this methodology as a future work by applying unsupervised learning algorithms such as clustering in order to observe the division of this data and conclude further society classification. This allows us to divide society into approximate intervals for a better understanding of this distribution.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

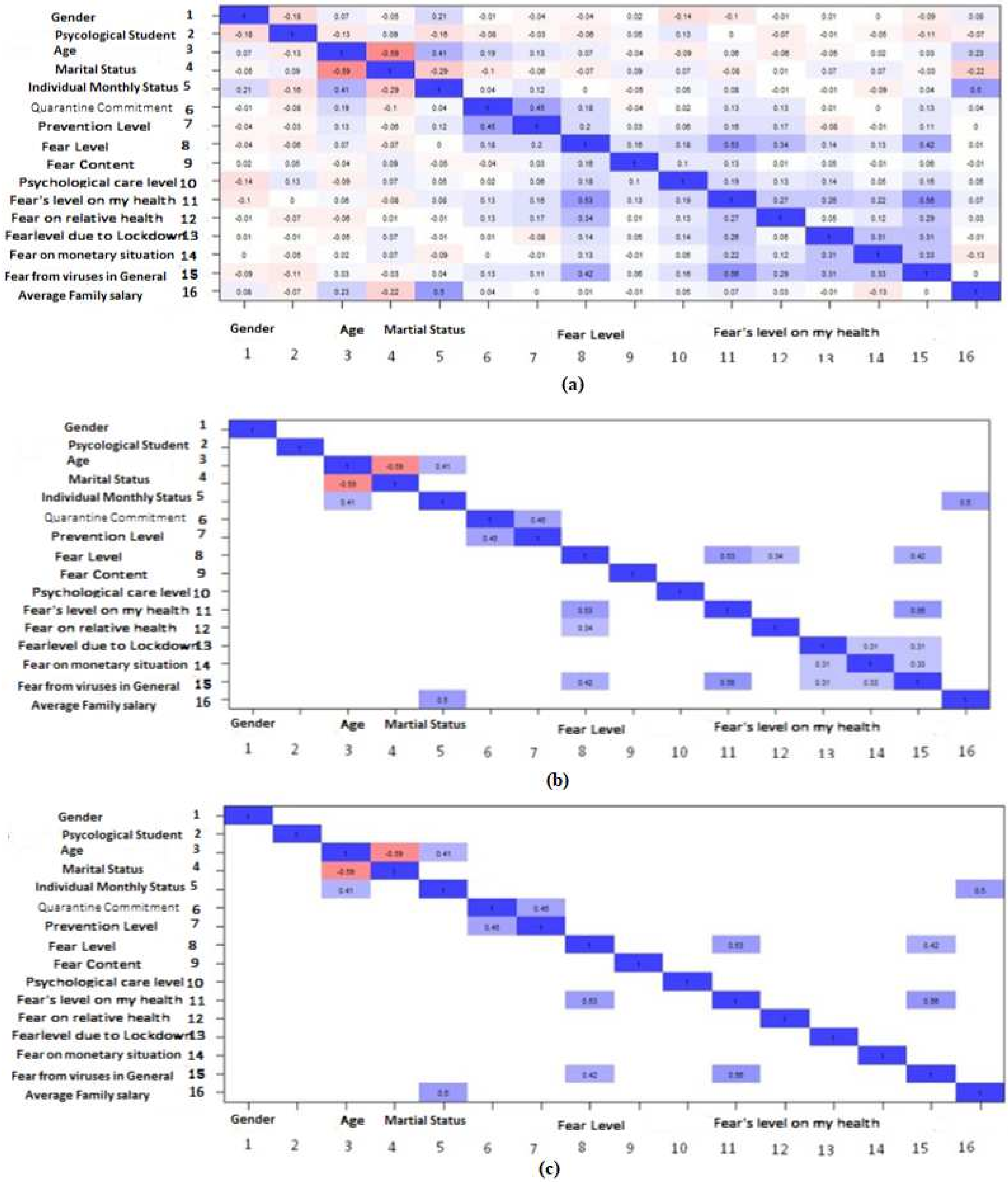

Fig. 6: (a) Confidence ϕ =0.6, variable support (σ). (b) Confidence ϕ =0.9, variable support (σ). (c) Variable Confidence (ϕ), Support σ =0.5.

Rules	Support	Confidence
{ I have a regular fear on my relatives' physical health } => { I have a severe fear from viruses in general}	0.3366	0.9613
{Female}=>{ I have a severe fear from viruses in general}	0.3211	0.9171
{ I have a regular fear on my relatives' physical health } => {I am Following up on this fear by communicating with a friend}	0.3230	0.8743
{Female}=>{I am Following up on this fear by communicating with a friend}	0.3636	0.9843
{Very poor} => {I collect Corona Info from TV;I collect Corona Info from social	0.3153	0.8232
{ I have a regular fear on my relatives' physical health } => {I collect Corona Info from TV;I collect Corona Info from social media }	0.3133	0.8182
{Female} =>{I collect Corona Info from TV;I collect Corona Info from social	0.3443	0.8990
{Very poor}=> { I have a severe fear on my monetary situation}	0.3501	0.8660
{ I have a regular fear on my relatives' physical health }=> { I have a severe fear on my monetary situation}	0.3462	0.8565
{Female}=> { I have a severe fear on my monetary situation}	0.3540	0.8756
{I Adhere on preventive measures from Corona in my daily habit}=> {Married}	0.3017	0.7256
{ I have a regular fear on my relatives' physical health }=>{Married}	0.3366	0.8093
{Female} =>{Married}	0.3578	0.8605
{Very poor}=> { I have a regular fear from viruses in general}	0.3540	0.8356
{ I have a regular fear on my relatives' physical health }=>{ I have a regular fear from viruses in general}	0.3404	0.8037
{Female}=>{ I have a regular fear from viruses in general}	0.3694	0.8721
{ I have a regular fear from being effected by Corona}=> {I have a regular fear on my physical health}	0.3075	0.7162
{Very poor} =>{I have a regular fear on my physical health}	0.3617	0.8423
{ I have a regular fear on my relatives' physical health }=>{I have a regular fear on my physical health}	0.3404	0.7928

Fig. 7: Association rules: Confidence ϕ=0.6, Support σ =0.3

References

[1] M. El Khoury-Malhame, R. Rizk, E. Joukayem, A. Rechdan, and T. Sawma, "The psychological impact of covid-19 in a socio-politically unstable environment: protective effects of sleep and gratitude in lebanese adults," BMC psychology, vol. 11, no. 1, p. 14, 2023.

[2] "Novel coronavirus (covid-19) deaths worldwide by country." https://www.statista.com/statistics/1093256/novel-coronavirus-2019ncov-deaths-worldwide-by-country/. Accessed: July 23, 2023.

[3] U. Nations, "United nations trade and development report 2022https://unctad.org/tdr2022," 2022. Accessed: July 23, 2023.

[4] A. Vysochyna, T. Vasylieva, O. Dluhopolskyi, M. Marczuk, D. Grytsyshen, V. Yunger, and A. Sulimierska, "Impact of coronavirus disease covid-19 on the relationship between healthcare expenditures and sustainable economic growth," International Journal of Environmental Research and Public Health, vol. 20, no. 4, 2023.

[5] Nature, "A briefing on developments in coronavirus research publishing.https://www.natureindex.com/news-blog/how-coronavirus-ischanging-research-practices-and-publishing," 2021.

[6] U. Ravens-Sieberer, A. Kaman, M. Erhart, J. Devine, R. Schlack, and C. Otto, "Impact of the covid-19 pandemic on quality of life and mental health in children and adolescents in germany," European Child & Adolescent Psychiatry, pp. 1–11, 2021.

[7] S. Li, Y. Wang, J. Xue, N. Zhao, and T. Zhu, "The impact of covid19 epidemic declaration on psychological consequences: a study on active weibo users," International journal of environmental research and public health, vol. 17, no. 6, p. 2032, 2020.

[8] A. A. H. Ahmadini, N. K. Adichwal, M. Z. Meetei, Y. S. Raghav, M. A. H. Ahmadini, A. Msmali, and N. Seth, "Knowledge, awareness and practices (KAP) about COVID-19 in jazan," Journal of Statistical Applications and Probability, vol. 10, pp. 487–497, July 2021.

[9] C. Stanley, A. Nkporbu, and P. Stanley, "Socio-economic implications of the coronavirus pandemic: Early review and projections, nigeria in focus," International Journal of Research Studies in Medical and Health Sciences, vol. 5, no. 4, pp. 40–49, 2020.

[10] M. Buheji, K. da Costa Cunha, G. Beka, B. Mavric, Y. De Souza, S. S. da Costa Silva, M. Hanafi, and T. C. Yein, "The extent of covid-19 pandemic socio-economic impact on global poverty. a global integrative multidisciplinary review," American Journal of Economics, vol. 10, no. 4, pp. 213–224, 2020.
[11] K. M. Antioch, "The economics of the covid-19 pandemic: economic evaluation of government mitigation and suppression policies, health system innovations, and models of care," Journal of Public Health, pp. 1–16, 2023.
[12] B. J. Middendorf, A. Faye, G. Middendorf, Z. P. Stewart, P. K. Jha, and P. V. Prasad, "Smallholder farmer perceptions about the impact of covid-19 on agriculture and livelihoods in senegal," Agricultural Systems, p. 103108, 2021.
[13] S. Gossling, D. Scott, and C. M. Hall, "Pandemics, tourism and global change: a rapid assessment of covid-19," Journal of Sustainable Tourism, vol. 29, no. 1, pp. 1–20, 2020.
[14] M. F. Bashir, M. Benjiang, and L. Shahzad, "A brief review of socioeconomic and environmental impact of covid-19," Air Quality, Atmosphere & Health, vol. 13, no. 12, pp. 1403–1409, 2020.
[15] T. Day, I.-C. C. Chang, C. K. L. Chung, W. E. Doolittle, J. Housel, and P. N. McDaniel, "The immediate impact of covid-19 on postsecondary teaching and learning," The Professional Geographer, vol. 73, no. 1, pp. 1– 13, 2021.
[16] W. Abd-Elhafiez and H. Amin, "The digital transformation effects in distance education in light of the epidemics (covid-19) in egypt," Information Sciences Letters, vol. 10, p. 141–152, Jan 2021.
[17] O. Y. M. Al-Rawi, W. S. Al-Dayyeni, and I. Reda, "COVID-19 impact on education and work in the kingdom of bahrain: Survey study," Information Science Letters, vol. 10, pp. 427–433, Sept. 2021.
[18] A. Martin, M. Markhvida, S. Hallegatte, and B. Walsh, "Socio-economic impacts of covid-19 on household consumption and poverty," Economics of disasters and climate change, vol. 4, no. 3, pp. 453–479, 2020.
[19] A. W. Bartik, M. Bertrand, Z. Cullen, E. L. Glaeser, M. Luca, and C. Stanton, "The impact of covid-19 on small business outcomes and expectations," Proceedings of the National Academy of Sciences, vol. 117, no. 30, pp. 17656–17666, 2020.
[20] J. Ratnasingam, A. Khoo, N. Jegathesan, L. C. Wei, H. Abd Latib, G. Thanasegaran, L. C. Liat, L. Y. Yi, K. Othman, and M. A. Amir, "How are small and medium enterprises in malaysia's furniture industry coping with covid-19 pandemic? early evidences from a survey and recommendations for policymakers," BioResources, vol. 15, no. 3, pp. 5951–5964, 2020.
[21] H. M. Bai, "The socio-economic implications of the coronavirus pandemic (covid-19): A review," ComFin Res. https://, 2020.
[22] S. Lahiri and M. Sinha, "A study of the socio-economic implications of the covid-19 pandemic," Australasian Accounting, Business and Finance Journal, vol. 15, no. 1, pp. 51–69, 2021.
[23] V. Leiva, E. Alcudia, J. Montano, and C. Castro, "An epidemiological analysis for assessing and evaluating covid-19 based on data analytics in latin american countries," Biology, vol. 12, no. 6, p. 887, 2023.