**International Journal of Healthcare Information Systems and Informatics**
Volume 17 • Issue 1

# Association Rules Extraction From the Coronavirus Disease 2019:
## Attributes on Morbidity and Mortality

Donald Douglas Atsa'am, University of the Free State, South Africa*
https://orcid.org/0000-0001-9687-8042

Ruth Wario, University of the Free State, South Africa
https://orcid.org/0000-0003-3733-3485

## ABSTRACT

This research was aimed to extract association rules on the morbidity and mortality of corona virus disease 2019 (COVID-19). The dataset has four attributes that determine morbidity and mortality; including Confirmed Cases, New Cases, Deaths, and New Deaths. The dataset was obtained as of 2nd April, 2020 from the WHO website and converted to transaction format. The Apriori algorithm was then deployed to extract association rules on these attributes. Six rules were extracted: Rule 1. {Deaths, NewDeaths}=>{NewCases}, Rule 2. {ConfCases, NewDeaths}=>{NewCases}, Rule 3. {ConfCases, Deaths}=>{NewCases}, Rule 4. {Deaths, NewCases}=>{NewDeaths}, Rule 5. {ConfCases, Deaths}=>{NewDeaths}, Rule 6. {ConfCases, NewCases}=>{NewDeaths}, with confidence 0.96, 0.96, 0.86, 0.66, 0.59, 0.51 respectively. These rules provide useful information that is vital on how to curtail further spread and deaths from the virus, both in areas where the pandemic is already ravaging and in areas yet to experience the outbreak.

## KEYWORDS

Association Rules Mining, Confirmed Cases, Corona Virus Disease, Data Mining, Deaths, New Cases, New Deaths

## INTRODUCTION

Association rules mining is one of the techniques in data mining that extracts interesting but hidden relationships among data objects in a dataset. The initial focus of association rules mining was to explore transaction databases for items frequently purchased together by customers (Mahmood, Shahbaz, & Guergachi, 2014). Modern research has successfully applied the topic in areas such as intrusion detection, telecommunications, disease diagnosis, and education (Mahmood *et al*., 2014;

DOI: 10.4018/IJHISI.302652
*Corresponding Author

Abdullah, Herawan, Ahmad, & Deris, 2011). According to Abdullah *et al*. (2011), two main steps are involved in association rules mining. In the first step, all frequent items are extracted from the transaction dataset. Frequent items are those that appear more than a specified number in the dataset. In the second step, common association rules are generated from the frequent items.

The coronavirus disease 2019 (COVID-19) is a viral infection caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV2) (Rothan & Byrareddy, 2020; Shereen, Khan, Kazmi, Bashir, & Siddique., 2020). According to Cortegiani, Ingoglia, Ippolito, Giarratano, and Einav (2020), Khan and Atangana (2020), and Shereen *et al* (2020), COVID-19 was first detected in the city of Wuhan, China in December, 2019. The World Health Organization (WHO) reports that as at 2nd April, 2020, COVID-19 cases had spread to 206 countries, territories and areas; infecting about 896,475 individuals. Out of this figure, a total of 45,525 individuals were reported to have died as a result of the disease (WHO, 2020). So far, there has been no approved vaccine against COVID-19 (Shereen *et. al.*, 2020) whose common symptoms include high body temperature, coughing and problems with breathing.

The objective of this study was to extract association rules from the morbidity and mortality attributes of the novel coronavirus disease 2019. The COVID-19 dataset has five variables that define the morbidity and mortality of the disease. These include *confirmed cases, confirmed new cases, confirmed deaths, confirmed new deaths,* and *number of days since last reported case*. The data, as presented by WHO (2020), does not provide information on how the disease variables on morbidity and mortality are associated. To address the gap, this study applies the association rules technique of data mining (Alola & Atsa'am, 2019; Atsa'am, 2020; Bodur & Atsa'am, 2019; Kantardzic, 2009) on the COVID-19 dataset to extract rules that show how the disease morbidity is associated with its mortality. The insights on how the various COVID-19 variables are associated in terms of antecedents and consequents should be useful in the global efforts to tame the pandemic.

## BACKGROUND

### Association Rules

Consider $X = \{x_1, x_2, ..., x_N\}$ as a set of $N$ distinct terms or items in a transaction dataset. Let *D* be a dataset of transactions, where each *T* is a subset of *X*. If *A* and *B* are sets of items then, an association rule is a notation of the form $\{A\} \Rightarrow \{B\}$ also written as $A \Rightarrow B$, where $A \subset X,\ B \subset X,\ \text{and}\ A \cap B = \varnothing$ (Abdullah, *et al*, 2011; Li L., Li Q., Wu, Ou, & Chen, 2018; Mahmood *et al*., 2014). A rule consists of two parts: the left-hand side (LHS) also called the antecedent; and the right-hand side (RHS), also called the consequent. Association rules mining is used for extracting interesting relationships among data objects in a dataset (Feng, Cho, Pedrycz, Fujita, & Herawan, 2016; Liu, Zhai, & Pedrycz, 2012). The technique is commonly used in market basket analysis for exploring the relationships between items purchased at supermarkets. In practice, an association rule will read for instance, "customers who purchase bread are 70% likely to also purchase butter". In medical terms, an established association rule will read "patients who exhibit cough, high temperature, breathing problems are 80% likely to have coronavirus disease". Real-world examples of how rules are typically represented are given in the Equation (1).

$$\left.\begin{array}{l}\text{Rule 1}: \{\text{bread}\} \Rightarrow \{\text{butter}\} \\ \text{Rule 2}: \{\text{cough, high temperature, breathing problems}\} \Rightarrow \{\text{COVID-19}\}\end{array}\right\} \quad (1)$$

The implication of Rule 1 in Equation (1) is that, the item on the RHS is frequently purchased alongside the item on the LHS. The Rule 2 implies that the item on the RHS is associated with the items on the LHS. That is, COVID-19 is the consequent of the antecedents: cough, high temperature, and breathing problem.

## Measures of Strength of Rules

There are three measures commonly used for evaluating the strength of a rule. These include: support, confidence, and lift (Hu & Chen, 2006).

### *Support*

Support is the ratio of the number of transactions that contain both the LHS and the RHS to the total number of transactions in the dataset. Referring to Equation (1), the support for Rule 1 is the ratio of all transactions where bread and butter were purchased together to the total number of transactions in the dataset. For Rule 2, the support is the ratio of those observations where the symptoms on the LHS were detected and COVID-19 was confirmed, versus the total number of observations in the dataset. The formula for computing support is given in Equation (2) (Ghafari & Tjortjis, 2019; Huang, Lu, & Duan, 2011).

$$Support = \frac{\text{Number of transactions with both LHS and RHS}}{\text{Total number of transactions}} = \frac{P(LHS \cap RHS)}{N} \tag{2}$$

### *Confidence*

The formula for computing confidence is given in Equation (3). Confidence measures the likelihood that the RHS will occur (or will be purchased) whenever the LHS occurs (or is purchased) (Ghafari & Tjortjis, 2019; Huang *et al.*, 2011).

$$Confidence = \frac{\text{Number of transactions containing both LHS and RHS}}{\text{Total number of transactions with LHS}} = \frac{P(LHS \cap RHS)}{P(LHS)} \tag{3}$$

### *Lift*

This is a measure of the likelihood that the item on the RHS will be purchased whenever the item(s) on the LHS is/are purchased, while considering the popularity of both items (Soysal, 2015). Possible values of Lift range from zero to infinity. When Lift $> 1$, it indicates that the LHS and RHS occur together more often than expected. When Lift $< 1$, it indicates that the chances of the LHS occurring together with the RHS are minimal. When Lift is close to 1, it shows that the LHS and the RHS appear together, almost often, as expected. Lift is computed using the formula in Equation (4) (Soysal, 2015).

$$Lift = \frac{\text{Confidence}}{\text{Expected Confidence}} = \frac{P(LHS \cap RHS)}{P(LHS).P(RHS)} \tag{4}$$

It is to be pointed out that in all measures of the strength of a rule, the higher the values, the stronger the rule. Consequently, when mining association rules, it is required to eliminate the rules with lower values of support, confidence or lift.

There are several algorithms for mining association rules from datasets, one of which is the Apriori algorithm. According to Li *et al.* (2018), the Apriori algorithm executes in the following steps: first, the frequent itemsets of length one are generated. This is repeated until all frequent itemsets have been identified. Next, all frequent itemsets of length k+1 are iteratively generated from those of length k. Then, all the candidate itemsets that contain subsets of length k which are not frequent are pruned. The algorithm then scans the dataset and counts the support of each candidate itemset. Lastly, infrequent candidate itemsets are eliminated, leaving only the frequent ones.

## METHODOLOGY

### The COVID-19 Data

The COVID-19 situation report consists of data generated from coronavirus cases across the world. The dataset is maintained by WHO, and the version used in this study contains 896, 475 cases as of 2nd April, 2020. The cases are across 206 countries, territories or areas (WHO, 2020). The dataset attributes are described below.

- **ConfCases:** This variable holds the total number of confirmed cases in a particular country, territory or area. According to WHO (2020), a confirmed case is someone whose laboratory test result indicates that they are infected with coronavirus, whether they show clinical signs and symptoms of the disease or not.
- **NewCases:** This variable holds the total number of confirmed new cases in a particular country, territory or area. For this research, this variable holds the total number of COVID-19 cases that were confirmed on 2nd April, 2020.
- **Deaths:** This holds the total number of deaths occasioned by COVID-19 in a particular country, territory or area.
- **NewDeaths:** This holds the total number of new deaths from COVID-19 in a particular country, territory or area on the current date. For this research, the variable reports the deaths that occurred on the 2nd April, 2020.
- **TransmissionClass:** This reports the mode of transmission of COVID-19 cases in a particular country, territory or area. There are five categories: community transmission, local transmission, imported case only, under investigation, and interrupted transmission. Where multiple modes of transmission have been reported, the WHO selects the category with the highest cases.
- **DaysLastCase:** For a country, territory or area, this variable holds the total number of days between the date a COVID-19 case was last confirmed and the current date when another case is confirmed.

### Data Preprocessing

The first preprocessing activity carried out was the selection of relevant attributes. Two COVID-19 attributes: *TransactionClass* and *DaysLastCase* have little or no relevance to the current study and were thus eliminated. Four attributes: *ConfCases, NewCases, Deaths, NewDeaths* were retained in the dataset. The COVID-19 data consists of values that range between zero and several thousands. The Apriori algorithm operates only on transaction dataset, which required that the COVID-19 data be converted to transaction format. In a transaction dataset, records are represented by item names. All items purchased in a transaction are enumerated in form of a record and blank spaces indicate when a particular item was not purchased. The COVID-19 data was transformed to transaction format by replacing all numeric values other than zero with the variable name of the corresponding attribute. Where data values were zeros, the zeros were removed and the space was left blank. Samples of 10 records from the COVID-19 dataset in transaction format were randomly selected to give insight on the structure of transaction data – see Table 1.

Blank spaces in Table 1 indicate that no COVID-19 incident in the corresponding attribute occurred in that country as of 2nd April, 2020. Where a COVID-19 incident was reported in a country as of 2nd April, 2020, the corresponding variable name was recorded as a data value. Consider the country, Vietnam, for instance. The COVID-19 transaction data shows that as of 2nd April, 2020; this country had at least one confirmed case, at least one new case, and no deaths or new deaths

Table 1. Sample of the COVID-19 dataset in transaction format

| Country/Territory | ConfCases | NewCases | Deaths | NewDeaths |
|---|---|---|---|---|
| China | ConfCases | NewCases | Deaths | NewDeaths |
| Japan | ConfCases | NewCases | Deaths | |
| Vietnam | ConfCases | NewCases | | |
| Cambodia | ConfCases | | | |
| Mongolia | ConfCases | NewCases | | |
| Nigeria | ConfCases | NewCases | Deaths | NewDeaths |
| Italy | ConfCases | NewCases | Deaths | NewDeaths |
| Spain | ConfCases | NewCases | Deaths | NewDeaths |
| United Kingdom | ConfCases | NewCases | Deaths | NewDeaths |
| United States of America | ConfCases | NewCases | Deaths | NewDeaths |

were recorded in that country. There are 206 total transactions in the dataset, and each transaction is a record of COVID-19 incidents in a country, territory or area as reported by (WHO, 2020).

The item frequency property of the dataset attributes was examined using the frequency plot in Figure 1.

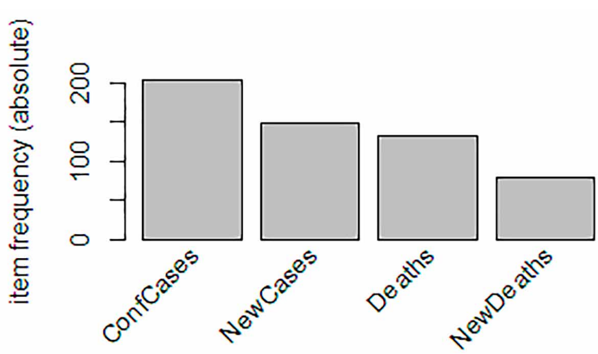
The Figure 1 shows that *ConfCases* is the most frequent item in the COVID-19 dataset, and the least frequent item is *NewDeaths*. This means that *ConfCases* occurs more often than other attributes while *NewDeaths* occurs less often in the observations.

## Association Rules Extraction

The preprocessed data (consisting of 206 records) in transaction format was uploaded to the R programming language environment, and the Apriori algorithm function was invoked. The R language codes used for association rules mining in this study are shown in Listing 1.

The first line of codes from Listing 1 invoked the Apriori algorithm on the data frame, *covid-19*. The minimum support and confidence were set to 0.01 and 0.4 respectively, and the results were returned through an object, *rules*. The second line of codes specified that the rules generated be sorted in decreasing order of confidence. The third line printed the extracted rules to the screen. The Apriori algorithm generated a total of 28 association rules from the COVID-19 dataset.

Figure 1. Item frequency plot



**Listing 1. R language codes**

```
#Rules extraction
> rules <- apriori(covid-19, parameter = list(sup = 0.01, conf = 0.4))
> rules <- sort(rules, by= "confidence", decreasing = T)
> inspect(rules)
```

## RESULTS

The minimum confidence of 0.5 was decided as the threshold to select the strongest rules. A total of 22 rules were discarded because of weakness or duplicity; and the six strongest rules, shown in Table 2, were retained.

From Table 2, the support of a rule gives a fraction (or percentage) of observations where that rule occurred in the dataset (Ghafari & Tjortjis, 2019; Huang *et al.,* 2011). Taking Rule 1 for instance, the support of 0.37 means 37% of the 206 observations (countries) in the COVID-19 dataset have *Deaths, NewDeaths, NewCases* occurring together. The confidence gives a percentage of assurance that an established rule is likely to be extracted from any other dataset aside from the experimental dataset (Atsa'am & Bodur, 2019). The confidence of 0.96 gives a 96% assurance that Rule 1 will always occur when mining association rules on any coronavirus disease dataset. The Lift values in all the six rules are greater than one. This shows that the LHS and the RHS of each rule will always occur together more than expected in any investigation (Soysal, 2015). The count is related to support, and it tells the number of observations where the given rule occurred. The count of 76 for Rule 1 shows that 76 (out of 206) observations (countries) have *Deaths, NewDeaths, NewCases* occurring together. This evaluates to 37%, the same with support for Rule 1.

## DISCUSSION

The implications of the extracted association rules in practical terms with respect to morbidity and mortality of COVID-19 are discussed below.

Rule 1: Any country, territory or area whose citizens have died in the past, or on the current date, from COVID-19, is 96% likely to witness more new cases of coronavirus infection among her citizens. This rule is supported by 37% of the countries where COVID-19 has been reported by (WHO, 2020).

Rule 2: Provided that the citizens of any country, territory or area have been infected by coronavirus and any of her citizens have died of COVID-19 on the current date, that country, territory or

**Table 2. Association rules of COVID-19 attributes**

| Rule No | LHS | RHS | | Support | Confidence | Lift | Count |
|---|---|---|---|---|---|---|---|
| [1] | {Deaths, NewDeaths} | => | {NewCases} | 0.37 | 0.96 | 1.3 | 76 |
| [2] | {ConfCases, NewDeaths} | => | {NewCases} | 0.37 | 0.96 | 1.3 | 76 |
| [3] | {ConfCases, Deaths} | => | {NewCases} | 0.56 | 0.86 | 1.2 | 115 |
| [4] | {Deaths, NewCases} | => | {NewDeaths} | 0.37 | 0.66 | 1.7 | 76 |
| [5] | {ConfCases, Deaths} | => | {NewDeaths} | 0.38 | 0.59 | 1.5 | 79 |
| [6] | {ConfCases, NewCases} | => | {NewDeaths} | 0.37 | 0.51 | 1.3 | 76 |

area is 96% likely to experience more new cases of coronavirus infection. A total of 76 out of 206 countries, territories or areas where COVID-19 has been reported by (WHO, 2020) have validated this rule.

Rule 3: So long as there is a confirmed case of COVID-19 in a country, territory or area, and there has been any death caused by COVID-19 in the past, that country, territory or area is 86% likely to have more new cases of COVID-19. The COVID-19 data from 115 out of 206 countries, territories or areas have confirmed this rule to be valid.

Rule 4: Any country, territory or area whose citizens have died from COVID-19, and whose citizens have been infected with the virus on the current date, is 66% likely to witness more new deaths from coronavirus infection among her citizens. This rule is supported by 37% of the countries where COVID-19 has been reported.

Rule 5: Provided that there is any confirmed case of COVID-19 in a country, territory or area, and any of her citizens have died from COVID-19, that country, territory or area is 59% likely to witness more new deaths from coronavirus infection among her citizens. This rule is supported by 38% of the countries experiencing COVID-19.

Rule 6: If any country, territory or area has citizens who are infected with coronavirus, and new cases of infection have been reported on the current date, then that country, territory or area is 51% likely to lose her citizens to death from COVID-19 as time goes on. The data from 76 out of 206 countries, territories or areas where COVID-19 has been reported is in support of this rule.

It should be observed that the RHS of Rules 4, 5 and 6 is *NewDeaths* and interestingly, their confidence values are not as high as those of Rules 1, 2 and 3 whose RHS is *NewCases*. This is an indication that the morbidity consequence of COVID-19 is statistically more significant than its mortality consequence. The association rules extracted in this study can be a valuable reference material in the efforts to contain the COVID-19 pandemic. As noted, the LHS of a rule is referred to as the antecedent while the RHS is called the consequent (Ghafari & Tjortjis, 2019). This implies that an unwanted consequent can be averted by taming the corresponding antecedents. These association rules can be a good reference point for health practitioners and policy makers involved in the fight against COVID-19. The Rule 1 shows that past deaths and new deaths resulting from coronavirus infection have a strong association with the possibility of new cases taking place within a geographic area. An effective way to curtail the number of new cases within an area is to prevent deaths and new deaths from the infection. Using the same argument for Rules 2 and 3, the number of new coronavirus cases within a geographic area has a strong association with past confirmed cases and past deaths or new deaths from the infection. Authorities in areas where COVID-19 has not yet been detected can learn from this information to control incidents of new cases to barest minimum provided such areas later on experience the disease. Rules 4, 5 and 6 provide valuable information that can be relied on by health practitioners to prevent new COVID-19 deaths within a given area. The three rules show that to reduce the number of new deaths, the number of deaths, confirmed cases, and new cases should be controlled. These rules are applicable to both areas where the pandemic is already ravaging and to the areas yet to experience the disease. While the former category of areas will cash on these findings to curtail further spread or deaths from the virus, the latter category will utilize these to put in place proactive measures to prevent spread and deaths.

## CONCLUSION

As of 2nd April, 2020, a total of 206 countries, territories and areas had their fair share of the coronavirus pandemic as reported in the WHO COVID-19 situation report – 73. This study deployed the Apriori algorithm to extract six association rules on the morbidity and mortality of the COVID-19. The COVID-19 data has two variables on morbidity: *ConfCases* and *NewCases*; and two variables on mortality: *Deaths* and *NewDeaths*. In three of the extracted rules, *NewCases* was the consequent variable while in the other three, *NewDeaths* was the consequent. Each rule shows the various combinations of variables that lead to the consequent. It is expected that the extracted rules will be referred to for public health policy formulation with respect to the fight against corona virus. The priority that each government attaches to the two consequents will determine the antecedents to be given more attention in the efforts to forestall or curtail future occurrences of the corresponding consequent. Geographic areas yet to experience the pandemic will utilize these rules to put in place proactive measures to prevent spread and deaths.

## REFERENCES

Abdullah, Z., Herawan, T., Ahmad, N., & Deris, M. M. (2011). Mining significant rules from educational data using critical relative support approach. *Procedia: Social and Behavioral Sciences*, *28*, 97–101. doi:10.1016/j.sbspro.2011.11.020

Alola, U. V., & Atsa'am, D. D. (2019). Measuring employees' psychological capital using data mining approach. *Journal of Public Affairs*, *2050*. Advance online publication. doi:10.1002/pa.2050

Atsa'am, D. D. (2020). Feature Selection Algorithm using Relative Odds for Data Mining Classification. In A. Haldorai, & A. Ramu (Eds.), Big Data Analytics for Sustainable Computing (pp. 81-106). Hersey, PA: IGI Global. doi:10.4018/978-1-5225-9750-6.ch005

Atsa'am, D. D., & Bodur, E. K. (2019). Knowledge mining on the association between psychological capital and educational qualifications among hospitality employees. *Current Issues in Tourism*. Advance online publication. doi:10.1080/13683500.2019.1597026

Bodur, E. K., & Atsa'am, D. D. (2019). Filter variable selection algorithm using risk ratios for dimensionality reduction of healthcare data for classification. *Processes (Basel, Switzerland)*, *7*(4), 222. doi:10.3390/pr7040222

Cortegiani, A., Ingoglia, G., Ippolito, M., Giarratano, A., & Einav, S. (2020). A systematic review on the efficacy and safety of chloroquine for the treatment of COVID-19. *Journal of Critical Care*, *57*, 279–283. Advance online publication. doi:10.1016/j.jcrc.2020.03.005 PMID:32173110

Feng, F., Cho, J., Pedrycz, W., Fujita, H., & Herawan, T. (2016). Soft set based association rule mining. *Knowledge-Based Systems*, *111*, 268–282. doi:10.1016/j.knosys.2016.08.020

Ghafari, S. M., & Tjortjis, C. (2019). A survey on association rules mining using heuristics. *Data Mining and Knowledge Discovery*, *9*(4), e1307. doi:10.1002/widm.1307

Hu, Y., & Chen, Y. (2006). Mining association rules with multiple minimum supports: A new mining algorithm and a support tuning mechanism. *Decision Support Systems*, *42*(1), 1–24. doi:10.1016/j.dss.2004.09.007 PMID:32287563

Huang, Z., Lu, X., & Duan, H. (2011). Mining association rules to support resource allocation in business process management. *Expert Systems with Applications*, *38*(8), 9483–9490. doi:10.1016/j.eswa.2011.01.146

Kantardzic, M. (2011). *Data Mining Concepts, Models, Methods, and Algorithms* (2nd ed.). John Wiley & Sons. doi:10.1002/9781118029145

Khan, M. A., & Atangana, A. (2020). Modeling the dynamics of novel coronavirus (2019-nCov) with fractional derivative. *Alexandria Engineering Journal*, *59*(4), 2379–2389. Advance online publication. doi:10.1016/j.aej.2020.02.033

Li, L., Li, Q., Wu, Y., Ou, Y., & Chen, D. (2018). Mining association rules based on deep pruning strategies. *Wireless Personal Communications*, *102*(3), 2157–2181. doi:10.1007/s11277-017-5169-0

Liu, X., Zhai, K., & Pedrycz, W. (2012). An improved association mining rules method. *Expert Systems with Applications*, *39*(1), 1362–1374. doi:10.1016/j.eswa.2011.08.018

Mahmood, S., Shahbaz, M., & Guergachi, A. (2014). Negative and positive association rules mining from text using frequent and infrequent itemsets. *TheScientificWorldJournal*, *973750*, 1–11. Advance online publication. doi:10.1155/2014/973750 PMID:24955429

Rothan, H. A., & Byrareddy, S. N. (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of Autoimmunity*, *109*, 102433. doi:10.1016/j.jaut.2020.102433 PMID:32113704

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*, *24*, 91–98. doi:10.1016/j.jare.2020.03.005 PMID:32257431

Soysal, Ö. M. (2015). Association rule mining with mostly associated sequential patterns. *Expert Systems with Applications*, *42*(5), 2582–2592. doi:10.1016/j.eswa.2014.10.049

World Health Organization. (2020, April 3). *Coronavirus disease (COVID-19) situation report –73* [Data set]. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7_6

*Donald Atsa'am is currently a Postdoctoral Research Fellow at the University of the Free State, South Africa. He has been a Computer Science lecturer with the University of Agriculture, Makurdi, Nigeria since 2012. He holds a Ph.D. in Applied Mathematics and Computer Science from the Eastern Mediterranean University, Famagusta, North Cyprus. His research interests are in Data Mining and Knowledge Discovery, Computational Intelligence, and Machine Learning. Donald is a Certified Information Systems Auditor with several years of experience in Systems Audit and Control.*

*Ruth Wario (PhD) is a Senior Lecturer at the Department of Computer Science and Informatics at the University of the Free State, South Africa. She is currently teaching and supervising both undergraduate and postgraduate students. She teach, programming, Software engineering, Human Computer Interaction, Database management and information management system. Her are of specialization is ICT4D and Human Computer Interaction.*