Abstract

The emergence of new technologies has modified the way businesses handle the information. Big data technology, which is best known for its capability to deliver competitive advantages, has been increasingly adopted by many organizations. A well-informed workforce of advanced technologies is fundamental for successful implication of any technology. Therefore, the technological understanding has become a fundamental expectation of corporate recruiters and further it determines successful career establishments. The purpose of present study is to assess the big data awareness among management undergraduates in Sri Lanka. A descriptive survey was performed and the correlation, multiple regression analysis and structural equation modeling performed the data analysis. The results disclosed that level of usefulness and the level of effectiveness on big data are significant components of big data awareness where level of experience found insignificant predictor of big data awareness. The findings implied the directions in which the future human resources should be developed.

Keywords: Big data, Sri Lanka, Structural Equation Modeling, Undergraduates
INTRODUCTION

Contemporary business entities function in a data economy where continual efforts are needed to track opportunities in the environment. A typical business operation of a day accumulates bulk data sets from different sources. These massive sets of data are called big data which are difficult to manage by traditional database management systems. According to Agrawal (2015), big data, as the name suggests, refers to large datasets that are challenging to store, share, search, visualize, and analyze. Here the orders of magnitude exceed conventional data processing and the largest of data warehouses. Usually data volume, velocity, and variety describe big data, but the unique attribute of big data is the manner in which the value is revealed. Big data has become one of the trending technologies that drastically changed the way which businesses handled their information. The dramatic growth & use of multi-technological innovations, the ubiquity of the internet & IT applications and the pervasiveness of smart mobile devices have increased real-time data collection and has exponentially led to utilization of big data in today’s businesses environment (Juniper Networks, 2012) However, due to the various forms and structures of these collected data, the tools and architectures needed for its real-time analysis may vary and could be a challenge for many organizations, especially those operating with limited IT budgets (Kalema & Mokgadi, 2017).

In the era of the fourth industrial revolution (Industry 4.0), big data has major impact on businesses, since the revolution of networks, platforms, people and digital technology have changed the determinants of firms’ innovation and competitiveness. An ongoing huge hype for big data has been gained from academics and professionals, since big data analytics leads to valuable knowledge and promotion of innovative activity of enterprises and organizations, transforming economies in local, national and international level. Nowadays, the tremendous increase of data through the Internet of Things- IoTs (continuous increase of connected devices, sensors and smartphones) has contributed to the rise of a “data-driven” era, where big data analytics are used in every sector (agriculture, health, energy & infrastructure, economics & insurance, sports, food and transportation) and every world economy (Vassakis et al., 2018).

The ways of how a company goes about using data to the best advantage and the means of transforming massive amounts of data into knowledge are some of the major concerns that companies are focused. A competent workforce thus will be a grounding requirement to reap the best out of new technologies.

As stated by Tennakoon & Lasanthika (2019), knowing what the big data is and their business value have become a dominant managerial competency of any capacity. Utilizing developing technologies and adapting to the 21st century would be possible by being informed about the technological developments and raising individuals who are able to comprehend the importance of technology in human and social life (Morgil et al.,
Enterprises in order to leverage data through big data analytics need human capital with high level of technical skills to use and exploit these systems in order to achieve exploitable knowledge for end users, mainly C-suite. People’s specific skills include statistics, big data mining, master visualization tools, business-oriented mindset and machine learning. These are required to get valuable insights from big data contributing in decision making procedure (McAfee et al., 2012).

As most of the organizations are moving with technological advancements and driven by big data capabilities, the understanding about new technologies, like big data will be a prime expectation of many recruiters. Its newness and relative complexity demand a proper awareness that supports to perform well with bulk data sets. Given the priority of managing the big data for organizational succession, big data awareness is a critical success factor of any organizational movement, the present study on big data awareness has appeared highly significant. Business organizations tend to seek candidates who have greater technological literacy. The undergraduates will step to the corporate world after successful completion of their studies. They should be essentially competent in managing big data for operational and strategic purposes of the organization (Zook et al., 2017). Further, companies will be able to achieve better results with the interns who are informed of new technologies. Their awareness of new technologies will assist them perform well with big data platforms.

But there is a shortage of empirical evidence on the significance of big data awareness among management undergraduates especially in Sri Lankan context. Realizing the level of big data awareness will guide higher education institutions in redesigning their curriculum so as to equip undergraduates with technological capabilities to address the demands of the corporate world.

The study primarily aimed at detecting the major contributor/s of the big awareness among the management undergraduates of Sri Lanka. Assessing the degree of each three dimensions of big data awareness is considered the specific objectives of the study.

**LITERATURE REVIEW**

Big Data is a concept often used when an organization’s existing traditional relational database and file systems processing capacities are exceeded in high transactional volumes, velocity responsiveness and or variety of data. The data are too big, move too fast, or don’t fit the structures of Rational Database Management System architectures. Scaling also becomes a problem. To gain value from these data, organizations must choose alternative ways to process the data sets (Goss and Veeramuthu, 2013).

Big data is not a technology itself. It refers to collections of data so large, varied and dynamic forms that they cannot be handled by conventional data processing technology.
There are four characteristics; 4Vs that define big data:

- **Volume** - The amount of data collected, managed and stored.
- **Velocity** - The speed at which it is received.
- **Variety** – The different types of data from multiple sources.
- **Value** – The insight of which the organization gain.

Processing big data is complex, due to its great variety, high velocity and extremely large volume (Kankanhalli et al., 2016). Yet, with advanced technologies this type of data can be combined and analyzed, revealing information that was hitherto virtually undiscoverable. Moreover, thanks to developments in computational, storage and analytical technologies, tools for handling and using this data are becoming ever more accessible (Bryant et al. 2008 & Hota et al. 2015).

Organizations could improve their decision-making thereby realizing their objectives if they use information embedded in big data sets at their disposal intelligently. As reported by Manyika et al., (2011), organizations may lose competitiveness if they fail to systematically analyze the available information needed for decision making. Big data can improve decision making and increase organizational efficiency & effectiveness, provided that organizations employ a variety of analytical tools and methods to make sense of the data (Joseph & Johnson, 2013).

The ability to manage big data, information and knowledge to gain competitive advantage and the importance of business analytics through big data has been well established in recent years. Organizations are investigating ways to efficiently and effectively collect & manage the data, information & knowledge which they are exposed to via various internal & external sources in the networked society.

According to Bussaban & Waraporn (2015), a growing number of bachelors in science and technology, particularly in computer science and mathematics conferred has increased in the last several years and the trend is continuing. This signifies strong workforces that can be deployed in both public and private sectors. Surprisingly, this is not valid as the skills trained are not fully maximized by most employers as they are not looking for those to work in solitude but rather trans-disciplinarily. In order to address this impending issue of hugely available but lack in quantitatively trained-skills human resources (Manyika, et al., 2011; Zorn, et al., 2014), changes in curriculum are apparent and indeed mandatory. This will ensure that undergraduates are ready for the practical data analysis work ahead using the real-world data. With the emergent rise of big data era brings about even more an immediate need of making aware undergraduates on new data technologies. Various leading associations stress the need to attract, train and retrain present and future batch of undergraduates (Rudin and Vahn, 2014; Fox & Hendler, 2014). Therefore, it is vital to explore how aware the undergraduates of big data are in terms of its usage, usefulness and effectiveness.
Development of Hypotheses

The present study proposes three different variables to evaluate the level of big data awareness among undergraduates. Those three dimensions are, Experience level of big data (Izhar & Shoid, 2016; Andreou et al., 2007; Kwon et al., 2014; Assuncao et al., 2015), Usefulness level of big data (Davenport and Dyche, 2013; Kwon et al., 2014; Erickson & Rothberg, 2014) and Effectiveness level of big data (Manyika et al., 2011; Grossman & Siegel, 2014; Galbraith, 2014; Esposito et al., 2015). First evaluates the level of personal experience on big data to see if undergraduates have experience on big data. This is important because some might experience big data issues that help them to come out with a strategy in managing big data. Second evaluates the usefulness level of big data and the last evaluates the effectiveness level of big data. The understanding of big data at these levels will measure a comprehensive understanding of big data. Further it guides the users when and where that understanding can be embedded to the organizational decision-making process.

Experience Level of Big Data

Understanding big data is very important and it can be gained through personal experience. Personal experience includes knowing the definition of big data, understanding the creation of big data and knowing the types of data in their organizations. Similarly, to be aware of investigating if their organizational data are part of big data itself. Personal experience on big data can be considered important to evaluate the level of big data awareness because experience can arm them with better strategy in managing big data (Izhar & Shoid, 2016). Even though some leading companies are actively adopting big data analytics to strengthen market competition and to open up new business opportunities, many organizations are still in the early stage of the adoption curve due to lack of understanding of and experience with big data (Kwon et al., 2014). Understanding big data can gain insights from the produced data and it is a key to competitive advantage (Assuncao et al., 2015). Accordingly, the study assumes that the experience level of big data as a key determinant of big data awareness.

H1: Experience level of big data determine the awareness on big data

Usefulness Level of Big Data

Evaluating the usefulness level of big data will assist the users to understand big data in daily business activity. It emphasizes how big data creates value in organizations by creating useful information and knowledge for better decision-making in supporting organization certain priorities (Davenport & Dyche, 2013; Grossman & Siegel, 2014; Manyika et al., 2011). The evaluation of usefulness will determine if big data matters in organizations from the aspect of tools analysis, improving data analysis and improving & optimizing decision-making. The most likely to initiate big data technologies are either
existing analytics groups or innovation groups within IT organizations (Davenport & Dyche, 2013). Level of usefulness includes if understanding big data can improve the process in managing large volumes of data in organizations (Izhar & Shoid, 2016). Thus, the usefulness is assumed to be a key component of big data awareness.

H2: Usefulness level of big data determine the awareness on big data

**Effectiveness Level of Big Data**

Effectiveness of big data emphasizes to which extent organizations have successfully managed big data in their organizations to extract insights (Galbraith, 2014). Additionally, it holds whether managing big data has enabled organizations to improve their competitive edge, decision-making, innovation and organizational capital (Andreou et al., 2007). More effective big data analyses can lead to more confident decision-making and better decisions can mean greater operational efficiencies in organizations (Davenport & Dyche, 2013). It will drive the process of knowledge creation and also understand their knowledge culture (Erickson & Rothberg, 2014). At the end, it will resolve the problem of knowledge pitfalls and create knowing culture at the organizations as part of organizational culture itself. Effectiveness of big data analytics is considered an imperative aspect to be further improved in order to increase the operating margin of both public and private enterprises, and represents the next frontier for their innovation, competition, and productivity (Esposito et al., 2015). Based on the presented thoughts, understanding the effectiveness level of big data is considered as a principle element of big data awareness.

H3: Effectiveness level of big data determine the awareness on big data

**CONCEPTUAL FRAMEWORK**

**Figure 1: Conceptual Framework**

Experience level of big data

Usefulness level of big data

Effectiveness level of big data

Awareness on Big Data

Source: Author Constructed
The presented variables are considered important in the research framework to drive better understanding of big data. The conceptual framework (provided in the Figure 1) is structured based on the discussion in the previous studies to evaluate the proposed variables; experience level, usefulness level and effectiveness level (Izhar & Shoid, 2016; Andreou et al., 2007; Assuncao et al., 2015; Davenport & Dyche, 2013; Erickson & Rothberg, 2014; Galbraith, 2014; Grossman & Siegel, 2014; Kwon et al., 2014; Manyika et al., 2011). Figure 1 depicted the presumed associations.

**METHODOLOGY**

A quantitative approach of the proposed research design was followed. Research method was a descriptive survey that adopted a self-administered questionnaire. Cross sectional primary data was gathered from respondents using an instrument of 24 items (Izhar & Shoid, 2016) of which 04 items devoted for demographic data, 03 for experience level of big data, 03 for usefulness level of big data, 09 items for effectiveness level of big data and 04 items for measuring awareness of big data by respondents. A five-point Likert scale measured the responses in which 1 stated “strongly disagree” and 5 denoted for “strongly agree”. Instrument satisfied all the validity and reliability requirements of all the study variables (Table 3). Sample was selected randomly from level four management undergraduates in the selected state universities of Sri Lanka. 390 questionnaires were distributed wherein 286 were returned (response rate = 73 %). Correlation, multiple regression analysis and structural equation modeling performed the data analysis with the aid of SPSS and SmartPLS 3 statistical software packages.

**RESULTS AND DISCUSSION**

**Table 1: Descriptive statistics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>N Statistic</th>
<th>Mean Statistic</th>
<th>Std. Dev Statistic</th>
<th>Skewness Statistic</th>
<th>Kurtosis Statistic</th>
<th>Std. E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience Level of Big Data</td>
<td>286</td>
<td>3.1926</td>
<td>1.01744</td>
<td>-.536</td>
<td>.279</td>
<td>.552</td>
</tr>
<tr>
<td>Usefulness Level of Big Data</td>
<td>286</td>
<td>3.2252</td>
<td>1.03678</td>
<td>-.550</td>
<td>.279</td>
<td>.552</td>
</tr>
<tr>
<td>Effectiveness Level of Big Data</td>
<td>286</td>
<td>3.2697</td>
<td>.98094</td>
<td>-.722</td>
<td>.279</td>
<td>.552</td>
</tr>
<tr>
<td>Awareness of Big Data</td>
<td>286</td>
<td>3.1239</td>
<td>1.10548</td>
<td>-.368</td>
<td>.279</td>
<td>.552</td>
</tr>
</tbody>
</table>

Source: Survey data (2019)
Demographics data showed that the majority of the respondents were female (76%), nearly 65% of respondents are having 01-02 years of work experience as interns in the areas of HRM, IT and Financial Management. Descriptive statistics evidenced an average level of overall understanding (Mean_Awareness = 3.1239) about the concept of big data. Additionally, experience level, useful level and effective level awareness too are in average level (Table 1). These results looked as if across the sample due to lower standard deviation values.

Results of correlation analysis depicted strong positive yet significant associations between independent variables and dependent variable. They are Experience level (r = 0.865, p = 0.000), Usefulness level (r = 0.873, p = 0.000) and effectiveness level (r = 0.878, p = 0.000).

Table 2 exhibits the outcome of multiple regression analysis. It explains the magnitude of Experience level, Usefulness level and Effectiveness level in predicting the big data awareness of respondents.

Table 2: Results of multiple regression analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R²</th>
<th>B</th>
<th>Sig.</th>
<th>Durbin-Watson</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.878</td>
<td>.770</td>
<td>.000*</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-</td>
<td></td>
<td>.003*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective_L</td>
<td>.878</td>
<td>.989</td>
<td>.000*</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.904b</td>
<td>.818</td>
<td>.000*</td>
<td>1.795</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-</td>
<td></td>
<td>.003*</td>
<td></td>
<td>.236</td>
<td>4.246</td>
</tr>
<tr>
<td>Effective_L</td>
<td>.878</td>
<td>.546</td>
<td>.000*</td>
<td>.236</td>
<td>4.246</td>
<td></td>
</tr>
<tr>
<td>Useful_L</td>
<td>.873</td>
<td>.480</td>
<td>.000*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:  
a. Predictors: (Constant), EFF_L  
b. Predictors: (Constant), EFF_L, USEF_L  
c. Dependent Variable: AWARE  
* coefficient is significant at .001 significance level

Source: Survey data (2019)

Model 2 in which the Effectiveness level and Usefulness level have identified as significant predictors of big data awareness was selected as the best fitted model. It is capable of predicting 81.8% of variation in the big data awareness. This is significant at
95% confidence level (p = 0.000). It satisfies all the assumptions of the multiple regression analysis including normality, autocorrelation, multicollinearity and heteroscedasticity (Table 1 and Table 2). Out of the hypothesized three predictors only two predictors were reported as significant. The experience level of big data awareness is not predicting the general awareness of big data in multiple regression analysis despite it was significant in correlation analysis.

Hence, the regression formulae can be derived as;

\[ Y = a + b_1 X_1 + b_2 X_2 + e \]  
\[ Y = -0.208 + 0.546 X_1 + 0.480 X_2 + e \]

Where;
- \( Y \) = Awareness Level of big data
- \( X_1 \) = Effectiveness level of big data
- \( X_2 \) = Usefulness level of data.

Thus,

\[ \text{Awareness of big data} = -0.208 + 0.546 \text{ (Effectiveness level of big data)} + 0.480 \text{ (Usefulness level of big data)} + e \]

Figure 2: Results of Path Analysis

Source: Survey data (2019)
Structural equation modeling is a multivariate statistical analysis technique that is used to analyze structural relationships. This technique is preferred by many researchers as it estimates the multiple and interrelated dependence in a single analysis. Having performed the multiple regression analysis, authors employed path analysis of structural equation modeling for further confirmation of the obtained results. The resulting path analysis along with the path coefficients and significance values of hypothesized paths are shown in the measurement model shown in Figure 2.

The overall effect of three exogenous variables (experience level of big data, usefulness level of big data, and effectiveness level of big data) found accounting 83.3% of the total variance of awareness of big data. This is statistically significant too (0.833, p = 0.000). However, the experience level of big data didn’t find statistically significant (0.215, p = 0.104) neither adequately describe the endogenous variable; awareness of big data. Similarly results of multiple regression analysis caused rejection of experience level of big data as a contributor of awareness of big data. This is not surprising as the respondents are not yet exposed to big data experiences at their courses and only aware of how useful and effective that it is. As such results confirm a significant explanatory power in usefulness level (0.331, p = 0.016) and effectiveness level of big data (0.404, p = 0.001). Both variables are capable of affecting the big data understanding at moderate level. Additionally, experience level appears to be linked with other two exogenous variables resembling that the mastery of a certain target always governs the understanding of usefulness and effectiveness of the target (Bandura, 1979). The results derived would not serve the purpose unless the validity and reliability of the measures are established. Table 3 presents the statistics on standards properties of the model and all the measures well satisfy the standard values.

Table 3: Statistics on the Properties of the Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cronbach Alpha</th>
<th>rho A</th>
<th>Composite Reliability</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience Level of Big Data</td>
<td>0.861</td>
<td>0.946</td>
<td>0.961</td>
<td>0.861</td>
</tr>
<tr>
<td>Usefulness Level of Big Data</td>
<td>0.883</td>
<td>0.935</td>
<td>0.958</td>
<td>0.883</td>
</tr>
<tr>
<td>Effectiveness Level of Big data</td>
<td>0.793</td>
<td>0.968</td>
<td>0.972</td>
<td>0.793</td>
</tr>
<tr>
<td>Awareness of Big data</td>
<td>0.854</td>
<td>0.944</td>
<td>0.959</td>
<td>0.854</td>
</tr>
</tbody>
</table>

Source: Survey data (2019)

The present study hypothesized three relationships to assess the awareness on big data among the management undergraduates in Sri Lanka. According to the statistical results,
the study found that the undergraduates are having an average level of awareness on big data and it is influenced by some factors. Moreover, the level of effectiveness and usefulness of big data found to have strong contribution toward big data awareness where level of experience on big data depicted an insignificant contribution. This implies that the students who are well informed about the usefulness and effectiveness of big data are having strong awareness on big data. As per the survey outcomes, the experience level on big data is not a significant predictor of big data awareness. This can be accepted as the undergraduates are not much involve with the practical big data operations and thereby the students are having low opportunity to enhance their big data awareness through experience.

The results of the present study consistent with some previous research works (Izhar & Shoid, 2016; Andreou et al., 2007; Assuncao et al., 2015; Davenport & Dyche, 2013; Erickson & Rothberg, 2014; Galbraith, 2014; Grossman & Siegel, 2014; Kwon et al., 2014; Manyika et al., 2011) while some inconsistencies are noted with some of the prevailing empirical evidences. Researchers suspect that the inconsistencies might be caused by the contextual dynamics and sample characteristics of the present study and would like to propose as research agendas for future studies.

CONCLUSION

Study aimed at determining the big data awareness of management undergraduates of Sri Lanka. Findings showed that the big data awareness is at moderate level and principally contributed by usefulness level understanding and effectiveness level understanding of big data. In addition, big data awareness of Sri Lankan undergraduates at present does not reflect any understanding about the experience level of it. It implies a lack of opportunities received by the undergraduates to practically expose to the big data experiences at their studies. They deemed to be having only the informative understanding about how useful and how effective that big data is. Yet, mastery experience is believed to be the most powerful source of knowledge and understanding about a certain phenomenon which will uplift the confidence of users to adopt it. Business organizations are expecting technologically-sought effective human resources who are flexible in adopting state-of-art technologies. As such the theoretical implications posit the importance of big data awareness and the power of each three dimensions of big data awareness in predicting the same. Further, the study generates empirical evidence relating to the existence of moderate level of big data awareness among the Sri Lankan undergraduates. The findings will add value to the existing body of knowledge in big data where scant literature found particularly in Sri Lankan context. Practical implications call for incorporating the big data knowledge into the curriculum to enhance the undergraduates’ awareness of it. In a way that will address the continued criticism of the corporate world about the mismatch of graduate’ attributes against the job requirements. Resultantly, higher education
institutions will be able to develop competent graduates who will best cater the industry demands. Implications of the study assist in bridging the knowledge gaps of new technologies where all the stakeholders get satisfied. Further the findings stress the importance of providing more opportunities to the undergraduates to gain experiences on new technologies like big data. The future researches would investigate the causal effect of big data awareness and managerial & organizational performance. Further, the future researchers are encouraged to test any possible mediating or moderating effects those strengthen the big data awareness.

LIMITATIONS

Authors located certain limitations of the present study mainly with respect to its methodology. In addition, the authors were unable to reach all the state universities and the survey covered only five state universities in Sri Lanka and also excluded the private universities.

REFERENCES


Juniper Networks (2012). Third Annual Mobile Threats Report, Juniper Networks Inc, USA.


