AN INVESTIGATION OF FACTORS AFFECTING EMPLOYABILITY OF BIG DATA PROFESSIONALS IN SRI LANKA; WITH SPECIAL REFERENCE TO LOGISTIC COMPANIES

Gunathilake, Lahiru¹, Weerasinghe, Vihanga¹,Gishanthi, Malsha¹,Kumarasinghe, Ovindi¹ and Bandara, Kelum¹

General Sir John Kotelawala Defence University¹,Kandawala Road, Dehiwala-Mount Lavinia (10390)

ABSTRACT

Big Data Analytics is swiftly growing and has revolutionized the field of business, through advanced analytics. Similarly, Sri Lanka is progressively embracing big data technology and the pioneering adopters include logistics companies. This emerging field has opened up many employment opportunities for big data professionals (BDP). However, Sri Lanka has encountered a shortage of BDP, amidst the significant growth in the field. Thus, this study analyze the factors that potentially impact the employability of BDP in the field of big data analytics, with the motive of finding solutions to reduce the skill shortage, which serves as the main objective of the research. The study was executed by analyzing qualitative and quantitative data collected through a questionnaire survey followed by a series of structured interviews. The questionnaire survey was distributed among 180 employees who are currently employed in the field of big data analytics, whereas, the structured interviews were carried out with 08 experts in the field. Based on the initial Exploratory Factor Analysis conducted, Education Factors, Skills and Competencies, and Job Market Factors were identified as the three main variables which affect the employability of BDP. Subsequently, a Thematic Analysis was carried out in order to investigate the impact of the aforementioned factors on the big data skill shortage, and to navigate possible remedies for it. Based on the data analysis conducted and results derived, it was depicted that the employability of BDP is directly related to Education Factors, Skills and Competencies and Job Market factors. As implications of the study it was revealed that certain educational and competency development factors should be enhanced in order to diminish the skill shortage of BDP.

KEYWORDS: Big Data Professionals, Demand, Skill shortage, Employability

Corresponding Author: Vihanga Weerasinghe, Email: vihanga.weerasinghe27@gmail.com

https://orcid.org/0000-0002-3296-6323

This is an open-access article licensed under a Creative Commons Attribution 4.0 International License (CC BY) allowing distribution and reproduction in any medium crediting the original author and source.
1. INTRODUCTION

Background of the study

In the present context, data has taken a new stance called “Big Data”. It is a pool of data that is massive in volume yet growing exponentially with time. LIRNE Asia, (2017) placed the development of Sri Lanka in the spotlight of big data, stressing the importance of up-to-date and accurate data, for a developing economy. The study relates to the field of logistics, in order to explore the shortage of BDP. This is mainly because it is one of the pioneering adopters of big data technology.

Figure 1 illustrates the latest statistics by World Economic Forum, regarding the technology utilization in the field of transportation and logistics. Big data analytics is ranked first among all the other technology enabled infrastructure, showcasing the potential of the field of logistics.

Figure 2 shows the investment on Big Data in the field of logistics. It distinguishes the forecasted increase in investments within the upcoming five-year period, showing the potential impact of Big Data on the field of logistics.

This study becomes unique and exclusive since it strives to investigate an aspect which has not grabbed significant attention from previous researchers. It is a highly technical field from the human perspective, by emphasizing the importance of human resources to carry out big data analytics. At the onset, the research study intends to scrutinize the reasons for the shortage of big data skill locally and globally. Initially, the study anticipates determining the most influential factors that affect the employability of BDP, through an Exploratory Factor Analysis.

Even though there is an increasing demand for BDP, Sri Lanka has encountered a shortage of supply of professionals to cater to this growing demand. World
Economic Forum (2019), Part (2010), Phillips (2017), and Rae (2018) examined that the job market has encountered a shortage of BDP. Therefore, the study strives to resolve the query, “What factors would affect the employability of BDP and what remedies could be undertaken to reduce the shortage of professionals, to reach the true potential of big data analytics?”

**Literature Review**

Ohlhorst 2013 expressed that big data is undertaken by many companies in the world as a main source of competitive advantage.

Even though the demand for BDP is rapidly increasing, the job market has indicated a significant gap in BDP, implying that the supply of professionals to the job market is poor, even though they are of high demand.

There is forecasted gap in-between supply and demand of BDP in US, for the year 2018. Based on that, the projected demand is much higher than the forecasted supply, resulting in a shortage in big data skill of 50%-60%.

Big data analytics is the scarcest skill in the global corporate field. As shown in the Figure 2.4, the market size of big data is expected to grow at a rapid rate.

Right human skill is critical in big data analytics (Dubey, et al., 2019; Wamba, et al., 2017). According to SHRM, 2016, 59% of organizations expect to elevate the job positions, which require the skill of data analysis, from 2017-2021.

**Figure 3: Expected growth of Big Data Market from 2011 to 2027 Source: Columbus, (2017)**

<table>
<thead>
<tr>
<th>Economy</th>
<th>Current DSA Workers</th>
<th>Projected DSA Workers</th>
<th>Percent Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malaysia</td>
<td>4,400 (2016)</td>
<td>20,000 (2020)</td>
<td>400%</td>
</tr>
<tr>
<td>The Philippines</td>
<td>147,420 (2016)</td>
<td>340,882 (2021)</td>
<td>131%</td>
</tr>
<tr>
<td>Singapore</td>
<td>9,200 (2015)</td>
<td>15,000 (2018)</td>
<td>61%</td>
</tr>
<tr>
<td>Canada</td>
<td>33,600 (2016)</td>
<td>43,300 (2020)</td>
<td>33%</td>
</tr>
<tr>
<td>United States</td>
<td>2,350,000 (2015)</td>
<td>2,720,000 (2027)</td>
<td>16%</td>
</tr>
</tbody>
</table>

**Figure 4: Demand of different skill levels for BDP Source: SHRM, (2016)**

Figure 4 interprets different skill levels required by employees. The analysis revealed that 60% of the organizations demand BDP with the ability to interpret and communicate results.

<table>
<thead>
<tr>
<th>Type</th>
<th>Name of the University/Institution</th>
<th>Academic program</th>
<th>Time Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>University of Moratuwa</td>
<td>Postgraduate certificate in data analysis and pattern recognition</td>
<td>1 year</td>
</tr>
<tr>
<td></td>
<td>University of Colombo School of Computing (USSC)</td>
<td>Master of Business Analytics</td>
<td>2 years</td>
</tr>
<tr>
<td></td>
<td>Informatics Institute of Technology (IIT)</td>
<td>BSc(Hons) Artificial Intelligence and Data Science</td>
<td>4 years</td>
</tr>
<tr>
<td></td>
<td>Sri Lanka Institute of Information Technology (SLIIT)</td>
<td>BSc (Hons) in Information Technology Specializing in Data Science</td>
<td>4 years</td>
</tr>
<tr>
<td>Private</td>
<td>National Institute of Business Management (NIBM)</td>
<td>BSc (Hons) Data Science</td>
<td>3 years</td>
</tr>
<tr>
<td></td>
<td>National Institute of Business Management (NIBM)</td>
<td>Advanced Diploma in Data Science</td>
<td>1 year</td>
</tr>
<tr>
<td></td>
<td>NSBM Green University Town</td>
<td>Professional Diploma in Data Science</td>
<td>1 year</td>
</tr>
</tbody>
</table>

factors anticipated by employers, when recruiting BDP. DASCA (2020) is a pioneering credentialing body for the data science profession. SAS (2020) is an international institute which offers certification to BDP, which is a value addition for them in career progression.

Time duration of higher educational qualifications play a significant role, since it gives a gist of the quality and capacity of the specific qualification.

The time durations of local undergraduate and postgraduate programmes relating to big data and advanced analytics are indicated in Table 1. This illustrates the duration of some selected degree and masters programmes that are related to big data and advanced analytics.

International educational platforms for big data analytics such as Pearson and Lytics Labs facilitate mainstream physical or virtual learning of various modules (Williamson, 2017).

IBM Data Science Professional Certificate is a platform which offers a vivid range of courses for professional BDP, which can be pursued independently even while engaging in employment (Widjaja, 2019). Oxford and Harvard Universities offer short term professional courses for BDP (Dhawan & Zanini, 2014).

However, the study of Ajah & Nweke, (2019) revealed that many organizations are not sufficiently equipped with knowledge and skill to implement big data analytics or to interpret the results of it. Therefore, it suggested the importance of building an organizational culture oriented on analytics, by bridging this skill and knowledge gap. Based on this, Harvey Nash/ KPMG CIO, (2020) revealed that 35% of the employers are anticipating to transform the workforce to polish their technology-related competencies.

According to the Figure 6, more than 76% of employers anticipate recruiting experienced BDP. The study of Park City Math Institute, (2016) stated that “Capstone projects” should be a mandatory component of the experience and internship programmes for Big Data employees.
Based on the Figure 7, as the experience grows, the remuneration of the big data employees increase.

Blake (2019) claimed that the field of data science is a very gratifying career which is increasingly relied upon by the society.

Aryal, et al. (2018), Cao (2017), Columbus (2017), and Wamba, et al. (2019) showed that embracing of big data technology in organizations require high level software like Hadoop, Apache pig and database management systems –NoSQL.

**Figure 8: Impact of COVID-19 pandemic on logistics and transportation companies**  

Figure 8 is a representation of different strategies that global logistics companies follow to adopt to the “New Normal”, post COVID-19.

### 2. METHODOLOGY

The research onion (Figure 9) which was developed by Saunders et al. (2009) explains the stages to be followed when developing a research strategy.

This study is conducted based on Pragmatism. Vallack (2010) explained that this philosophy is ideally used for research studies conducted based on a mixed method.

An inductive approach is selected based on the layout and execution of the research study.

The study involves the use of structured interviews and questionnaire surveys as its research strategies to collect qualitative and quantitative data. Thus, this study adopts a mixed method approach. Further, the study uses a cross sectional approach, where the information is gathered at a particular point of time.

**Population**

The target population for the series of structured interviews is the experts in the field of Big Data, and the respondents for the questionnaire survey are employees of selected logistic companies.

**Sample**

In order to collect data for the questionnaire survey, sample size of 180 operational and management level respondents in the field of technology were requested to fill in a questionnaire. The sample size was determined by the Morgan & Krejcie (1970) table, based on the method of simple random sampling under probability sampling. Structured interviews were conducted with selected Industry experts in the field of Big Data, and the sample size was decided by locating the saturation point after investigating the responses.

**Exploratory Factor Analysis**

This research is a very unique study and it explores an area which was not overlooked by many prior researchers. Thus, the researchers lacked firm theory and substantial models to support the conceptual framework in order to develop a hypothesis.
Therefore, an Exploratory Factor Analysis was conducted at the beginning, in order to determine the factors which affected the employment of BDP.

Determining the factors

The researchers initially determined certain indicators, that would possibly be affecting the employment of professionals in the field of big data analytics. Those indicators were chosen randomly, based on the literature survey. The aforementioned indicators include Competency of employees, Academic Knowledge, Higher Educational Qualifications, Remuneration, Experience, Soft Skills, Managerial Skills, Orientation of Qualifications, Professional Qualifications, Existing professionals in the field, Recognition, Infrastructure, Accreditations, Professional Networks and Time Duration.

Data Collection

Based on the indicators identified by the researchers, a questionnaire was developed and circulated and the researchers considered the responses of 180 operational and managerial level employees, in order to conduct the Exploratory Factor Analysis.

Preparation of data for analysis

Removing outliers

Based on the data collected, there were four potential outliers, and they were excluded from the dataset as shown in the Table 2.

Table 2: Data Screening

<table>
<thead>
<tr>
<th>Source: Sample Survey (2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire responses collected</td>
</tr>
<tr>
<td>Questionnaires discarded</td>
</tr>
<tr>
<td>Questionnaires considered</td>
</tr>
<tr>
<td>Outliers Removed</td>
</tr>
<tr>
<td>Questionnaires utilized</td>
</tr>
</tbody>
</table>

Table3: Guidelines for KMO Values

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td>Average</td>
<td>0.5 – 0.6</td>
</tr>
<tr>
<td>Acceptable</td>
<td>0.6 – 0.7</td>
</tr>
<tr>
<td>Good</td>
<td>0.7 - 0.8</td>
</tr>
<tr>
<td>Excellent</td>
<td>&gt;0.8</td>
</tr>
</tbody>
</table>

Table4: KMO Bartlett’s Test

<table>
<thead>
<tr>
<th>Source: Sample Survey (2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</td>
</tr>
<tr>
<td>Bartlett’s Test of Sphericity</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

According to Table 4, outputs indicate that KMO sampling adequacy value is 0.693 which is considered as an acceptable value, according to Hutcheson & Sofroniou (1999). The matrix can be ruled out if the Sig. value of the test is less than 0.005 (Field, 2000; Pallant, 2013). Therefore, since the sig. value in Bartlett’s Test of Sphericity is less than 0.005, the data set is adequately sampled.

Principle Component Analysis

A Principle Component Analysis (PCA) was conducted in order to distinguish the factors affecting employment of BDP. According to the initial analysis conducted, Higher Educational Qualifications, Academic Knowledge and Remuneration were identified as three doubtful indicators. This was because, their communality value was below 0.3 and the value in the component matrix was less than 0.5. If the communality value is less than 0.3, then it means that only less than 30% of the variance in this indicator shares a common origin with others. Therefore, those indicators should be excluded from the analysis (Hadi, et al., 2016). The component matrix displays the factor loadings without rotating the variables. If it contains any indicator less than 0.5, then the impact of that indicator to that specific variable is considered to be negligible.
After filtering those three indicators, the same Principal Component Test was carried out. Then the researchers considered the output of the “Total Variance Explained” in Table 4.4.

The parallel analysis in the Table 6 shows how the number of variables were determined to carry out factor extractions. In this process, the Eigenvalues obtained from PCA are compared with the Eigenvalues generated by Patil et al. (2017). The two values are compared in a way that if the Eigenvalue generated from PCA is greater than that of the parallel analysis, then the indicator is accepted (Horn, 1965).

After determining the number of independent variables for factor extractions, then the factors are rotated for further analysis. The main motive behind factor rotation is to align them in a way which makes it more convenient for interpretations.

Based on the Rotated Component Matrix in Table 7, the researchers loaded the factors to three main variables. The factor loadings were done so that the factors with a value more than 0.5 are grouped into categories based on their orientation.

The researchers labelled the factors based on the composition of indicators in them. This was based to develop the conceptual framework for the study.

Normality

Table 9 shows the normality measures of the data set of the research study.

Rose et al. (2015) mentioned that if the standard error of skewness and Kurtosis are within the range of +1.96 and -1.96, then the data set is normal. Since the
standard errors of the data set shown in Table 4.8, are in between this range, and it can be concluded that it is normal.

**Table 7: Rotated Component Matrix (Source: Sample Survey 2020)**

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing professionals</td>
<td>.786</td>
<td>.185</td>
<td></td>
</tr>
<tr>
<td>Recognition in the local job market compared to foreign job market</td>
<td>.756</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adoption of new infrastructure</td>
<td>.686</td>
<td>.118</td>
<td></td>
</tr>
<tr>
<td>Effect of industry experience</td>
<td>.643</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional networks</td>
<td>.596</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence of sufficient local qualifications</td>
<td>.827</td>
<td>.230</td>
<td></td>
</tr>
<tr>
<td>Crash courses to be completed in less time</td>
<td>.804</td>
<td>.204</td>
<td></td>
</tr>
<tr>
<td>Accreditation body provide guidance</td>
<td>.131</td>
<td>.636</td>
<td>.215</td>
</tr>
<tr>
<td>Professional qualifications</td>
<td>.611</td>
<td>.126</td>
<td></td>
</tr>
<tr>
<td>Effect of business and managerial skills</td>
<td>.130</td>
<td>.827</td>
<td></td>
</tr>
<tr>
<td>Satisfaction level of local graduates compared to foreign graduates</td>
<td>.801</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of soft skills</td>
<td>.148</td>
<td>.691</td>
<td></td>
</tr>
</tbody>
</table>

**Table 8: Labelling Factors**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Labels</th>
<th>Indicators</th>
<th>Factor Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 Job Market Factors</td>
<td>Existing Professionals</td>
<td>.786</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recognition</td>
<td>.756</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Infrastructure</td>
<td>.686</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>.643</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Professional Networks</td>
<td>.596</td>
<td></td>
</tr>
<tr>
<td>02 Educational Factors</td>
<td>Orientation of Qualifications</td>
<td>.827</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time Duration</td>
<td>.804</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accreditations</td>
<td>.636</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Professional Qualifications</td>
<td>.611</td>
<td></td>
</tr>
<tr>
<td>03 Skills and Competencies</td>
<td>Managerial Skills</td>
<td>.827</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Competency of Graduates</td>
<td>.801</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soft Skills</td>
<td>.691</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Labelling Factors

Source: Sample Survey (2020)

The researchers utilized Expert Validity technique, since the relationship between variables are yet unknown until the Exploratory Factor Analysis is conducted.

Reliability test was done by getting the Cronbach’s alpha value in SPSS. Bernstein (1994) confirmed that the standard value for Cronbach’s alpha could be more than 0.6, which was previously recommended by Bagozzi R.P. (1988). The Table 4.9 represents the reliability values of each variable utilized in this research study along with the number of indicators in each variable.

**Table 9: Skewness and Kurtosis (Source: Sample Survey -2020)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Skewness</th>
<th>Std. Error of Skewness</th>
<th>Kurtosis</th>
<th>Std. Error of Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education Factors</td>
<td>-.187</td>
<td>-.613</td>
<td>-.403</td>
<td></td>
</tr>
<tr>
<td>Skills and Competencies</td>
<td>.198</td>
<td>.198</td>
<td>.198</td>
<td></td>
</tr>
<tr>
<td>Job Market Factors</td>
<td>-.208</td>
<td>.119</td>
<td>-.267</td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td>.394</td>
<td>.394</td>
<td>.394</td>
<td></td>
</tr>
</tbody>
</table>

**Testing for Validity and Reliability**

The researchers utilized Expert Validity technique, since the relationship between variables are yet unknown until the Exploratory Factor Analysis is conducted.

Reliability test was done by getting the Cronbach’s alpha value in SPSS. Bernstein (1994) confirmed that the standard value for Cronbach’s alpha could be more than 0.6, which was previously recommended by Bagozzi R.P. (1988). The Table 4.9 represents the reliability values of each variable utilized in this research study along with the number of indicators in each variable.

**Table 10: Reliability for each variable**

(Source: Sample Survey -2020)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cronbach's Alpha</th>
<th>No of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational Factors</td>
<td>.715</td>
<td>4</td>
</tr>
<tr>
<td>Skills and Competencies</td>
<td>.620</td>
<td>4</td>
</tr>
<tr>
<td>Job Market Factors</td>
<td>.718</td>
<td>6</td>
</tr>
<tr>
<td>Employment</td>
<td>.644</td>
<td>3</td>
</tr>
</tbody>
</table>
Testing for Multicollinearity

Table 11: Multicollinearity (Source: Sample Survey -2020)

<table>
<thead>
<tr>
<th>Factors</th>
<th>Tolerance</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education Factors</td>
<td>.961</td>
<td>1.040</td>
</tr>
<tr>
<td>Skills and Competencies</td>
<td>.994</td>
<td>1.007</td>
</tr>
<tr>
<td>Job Market Factors</td>
<td>.899</td>
<td>1.112</td>
</tr>
</tbody>
</table>

If the tolerance value of the variables exceeds “one”, then there is no multicollinearity between the variables. However, if this value equals to “zero”, then the variables show perfect multicollinearity. Therefore, based on the Table 2.10, the variables considered in the study are proven to have no multicollinearity. The acceptable range of VIF value is between 10 and 0.1 (Field, 2005). The variables of the study abide by this rule as well. Hence, it shows that there is no multicollinearity prevailing among the variables.

Descriptive Statistics

Demographic profile

The 150 respondents were categorized into five groups under different age levels as 18-25, 26-35, 36-45, 46-55, 56 and above. Among them, the age group between 26-35 represented the majority (44%) of respondents. Significantly, there are no responses recorded from 56 and above age category. This gives an indication of the novelty of the field of big data. It shows the increased attention and attraction towards the field by younger generation when compared to older generation. Since only a very few respondents above the age of 46 have responded to the questionnaire, it could be assumed that the field of big data is not much embraced by the employees belonging to that age limit (Figure 10).

When considering about the Highest level of education, only 7% of the respondents belonged to the category of PhD. The highest percentage of respondents are in the group who have completed their Bachelor’s Degree (55%). As Figure 11 demonstrates, all the respondents have acquired more qualifications than Advanced Level.

Researchers further analyzed whether the respondents have acquired their highest educational qualifications from a local or a foreign University/institution. Among the 150 respondents, majority have acquired their highest educational qualification from foreign Universities/institutions, and it amounts to 52%. There are 48% employees who have acquired their highest educational qualification locally, which is comparatively lesser than graduates from foreign Universities. This demarcates the lack of higher educational platforms for BDP locally. The Figure 12 shows the orientation of the educational qualifications of respondents, based on the country and region.

Based on the Figure 13, majority of the respondents were data analysts, which summed up to a percentage...
of 33.22%. Both data scientists and data warehouse managers represented 17.11% each, which were the second highest. Meanwhile, the respondents include only a very less number of database managers.

In order to analyze the most commonly utilized software in Sri Lanka, five main types of software which are commonly used by BDP are considered and the respondents rated them based on their level of utilization. According to the Figure 14, most professionals heavily utilize Apache Spark, while Hadoop is ranked second based on high usage. Meanwhile, Hadoop and Apache spark are utilized mostly in the moderate usage category as well. Cassandra, R and Tableau are rated as low usage software by the respondents.

Researchers intended to analyze the impact of the COVID-19 pandemic to the big data employees. Based on their responses, most companies have initiated teleworking platforms due to the “New Normal” culture in the business sector. This was mentioned by 47% of the respondents, which summed up to be the highest. 36% of the employees have mentioned that they have been introduced with flexible working hours. 13% of the respondents have commented that all the traditional big data related platforms were transferred to the cloud. Based on the Figure 15, only a few number of employees experienced salary cut downs. This implies that COVID-19 has very slightly impacted the field in a negative manner.

Thematic Analysis

The research study employs thematic analysis to review and analyze qualitative data collected through a series of structured questionnaires.

Initial Reading of Texts

Braun & Clarke (2012) explained that this initial step is very vital to understand the content and the basic idea derived from various aspects.

Coding the Texts after Repeated Reading

In this phase, researchers generated 108 basic codes after the comprehensive analysis of the responses.

Generating Themes through Codes

The concept behind the generation of themes is the process of consolidation of the codes into like groups (Attride-Stirling, 2001; Braun & Clarke, 2006;
Lincoln & Guba, 1985). The final result obtained by the researchers included three main themes called Educational Factors, Skills and Competencies and Job Market Factors. Apart from that, another two codes; namely, Importance of Big Data in Logistics and Impact of COVID-19 on the field, were separately considered by the researchers, based on their significance to the field.

Presentation of data was in such a way that the researcher coded initially along with the themes generated through them.

**Thematic Network**

The thematic network shown in Figure 16 was developed by the researchers by investigating the codes developed through the interview texts.

**Data Analysis**

The data analysis is conducted by analyzing the themes and codes generated by the researcher.

**Importance of Big Data in the field of Logistics**

All the respondents agree that big data is perfectly compatible with the field of logistics. They placed the value of big data for the field of logistics in the “High” category, since it is a unique yet a valuable integration of two fields.

**Educational Factors**

**Accreditations**

Many have agreed that there are no recognized accreditation bodies locally. However, 63% of the respondents have commented that there are ample recognized accreditation bodies for big data professionals globally. Adding to this, Respondent eight has stated that there are recognized accreditation bodies for BDP, such as DASCA, IOA and CAP by Informs.

**Time Duration**

Half of the respondents commented that there are crash courses for BDP. Justifying this, Respondent seven has said that there are online platforms like Coursera. Meanwhile, respondent eight has also said that there are online crash courses offered by IBM and Google. Conversely, the remaining 50% of the respondents have commented that there are no crash courses for the employees in the field of big data analytics.

**Orientation of Qualifications**

Respondent five and six have commented that the standard of local graduates who get graduated from local Universities is high when compared to foreign graduates. However, majority of the respondents have stated that the quality and standard of foreign qualifications are high when compared to local qualifications.

**Professional Qualifications**

Three Respondents have commented that
professional qualifications are scarce in the local context. Majority of the respondents have agreed that there are ample professional qualifications available for BDP internationally. Respondent one has mentioned Udemy and Coursera as examples, in order to justify her point of view. Respondent two, three and six have commonly mentioned AWS as a recognized professional qualification. Respondent seven has given multiple examples such as Cloudera, Hortonworks, Elasticsearch, AWS, Azure, Cloud, Datadog and Snowflakes.

Skills and Competencies

Soft Skills

All the respondents have commonly expressed that soft skills are very important for BDP. They have specifically said that Communication is the most important skill, since it enables the professionals to express their findings to the top management and to the clients.

Managerial Skills

Majority of the respondents have agreed that managerial skills are critical for BDP. Justifying this, many respondents have collectively stated that people handling, time management, cost management, critical thinking and project management skills are vital for BDP. However, the respondents four, seven and eight have mentioned that it is not very critical for big data employees to possess managerial skills, since they engage in a technical role rather than in a managerial role.

Competency of Graduates

Four respondents have stated that many companies implement ample competency development programmes for employees in the field of big data. Respondent seven has interestingly mentioned that even though his company extends training and development programmes for BDP, they mostly expect the employees to be self-taught. Respondents one, four and six mentioned that there are only a smaller number of competency development programmes for BDP in local companies and they only cater to the specific requirements of employees. Also, Respondent six has stated that the company he is employed in has newly initiated training and development programmes for BDP. However, respondent two has strongly mentioned that there are no competency development programmes for BDP as at now.

Job Market Factors

Experience

All the respondents have commonly agreed that level of experience is a major factor affecting the employment of BDP. Respondents one and five have mentioned the importance of experience to drive performance and career progression of an employee.

Recognition

The researchers identified that the opinion of the majority of the respondents was that big data analytics as a profession is not yet recognized in Sri Lanka. According to Respondents one, three, four, five and seven, the field is still in the emerging stage, which is the main cause for the lack of recognition. Respondent two has captivatingly mentioned big data analytics as a “Surprise Field”.

Existing Professionals

Based on the responses received, respondents one, six and seven have mentioned that the number of existing professionals in the field locally are very less. Meanwhile, respondent two has mentioned that the existing professionals have so much to improve when compared to foreign professionals.

Infrastructure

Majority of the respondents have stated that most companies in Sri Lanka constantly update their infrastructure related to big data analytics. Respondent six has stated that their company conducts a lot of research on latest developments of infrastructure. However, respondent three has given
a contrary opinion saying that local companies do not regularly update their infrastructure.

**Professional Networks**

All the respondents have commented that professional networks are of utmost importance to the employees in the field of big data.

3. DISCUSSION

The conceptual framework (Figure 3.2) was developed by the researchers, based on the Exploratory Factor Analysis conducted at the onset of the research study. The relationship and correlation between the Independent Variables and the Dependent Variable was revealed through that, and this model was developed as an outcome of the same.

![Figure 17: Conceptual Framework Source: Developed by the researcher](image)

Considering this, three main hypotheses were developed.

- **H1a**: Educational Factors impact the Employability of BDP
- **H1b**: Skills and Competencies impact the Employability of BDP.
- **H1c**: Job Market Factors impact the Employability of BDP.

The outcomes of both the Exploratory Factor Analysis and the Thematic Analysis displayed a significant similarity in spite of the slight differences between responses.

**H1a: Educational Factors impact the Employability of BDP**

DASCA, 2020; the standards body ensures that all its accredited institutions are hiring destinations that are preferred by most organizations. Thus, accreditations of local and foreign bodies have a high impact on the employability of BDP.

Based on the research findings, there are many contemporary crash courses in big data analytics as well; namely, Coursera, 2020; EDX, 2020.

Even though there is only a less number of local qualifications on big data analytics, there are many foreign qualifications. SAS in collaboration with Birmingham City University has launched a programme called SAS Student Academy. Many local BDP declare that there are not much recognized higher educational platforms for big data analytics in Sri Lanka, which has to become a major concern in order to increase the number of professionals in the job market.

The experts in the field of big data mention that platforms such as AWS, 2020 facilitate professional education. Majority of the respondents have stated that they constantly feel the need of a standard professional qualification when working in a corporate setting, in order to update their knowledge and to climb up the corporate ladder. However, a concern is raised on the lack of professional qualifications locally, which might be an influential factor contributing to the current big data skill shortage.

**H1b: Skills and Competencies impact the Employment of BDP**

Interactive disciplines of Big Data Analysts should embrace soft skills such as critical thinking, creative thinking and communication (Song, 2016). All the respondents of the questionnaire survey commented that soft skills are of utmost importance to big data employees as well.
An investigation of factors affecting employability of big data professionals in Sri Lanka; with special reference to logistic companies

SAS, The Tech partnership, 2014 specified that the employers are interested in potential BDP with interpersonal, management and business insights. Majority of the respondents, along with a majority of the interviewers have agreed that managerial skills are important for employees in the field of big data.

Royster, 2013 stated that rounded up knowledge about the industry that the BDP are employed in will uplift their contribution to the sector. In order to do so, they should possess many competencies polished with traditional and up-to-date proficiencies. Based on the responses of the questionnaire survey it was concluded that the competency level of local BDP is satisfactory when compared with foreign professionals.

**H1:** Job Market Factors impact the Employment of BDP

In the Sri Lankan context, JKH, 2020 and PickMe, 2016 have specified minimum two years of experience in the field, for an employee to be recruited as a big data employee. All the industry experts have commented that experience is very important when employing BDP in an organization. Similarly, big data employees mention that the experience, skills and competencies that they have acquired through past experience have immensely helped them in performing their current jobs.

Certain employees find their career path based on the prestige of the field. Similarly, BDP also consider the recognition of the profession when engaging in employment. According to Hopkins & Hawking, 2018, big data analytics is growing in recognition in the global job market with its rapid evolution and potential strategic competitive inferences. Even though big data analytics is high in recognition in the international job market, it is not so in Sri Lanka. Industry experts comment that this is mainly because the industry is still in early stages of emergence. They further explain that only the professionals in the field are aware of the term “Big Data Professionals”, and in layman terms they are referred to as “Computer Engineers”.

Carillo, et al., 2019; Carillo, 2017; Intezari & Gressel, 2017; and Murawski & Bick, 2017 stressed the importance of on-the-job training, career guidance and continuous professional development programmes to develop analytical and technological skills. Similarly, Wickramasinghe, 2017 stressed on “Retrain to retain”, to overcome the employee shortfall by training the existing workforce to possess futuristic yet vital data analytic skills.

Many experts mention that even though the existing BDP in Sri Lanka are well-knowledge and talented, there is a scarcity of professionals in order to cater to the growing demand. However, they further mention that when compared to the professionals in developed countries, local BDP have so much to develop. When considering about the age of professionals, it is very visible that most employees in the field are young and energetic, but not very mature in age.

4. CONCLUSION

The research study identified that the employability of BDP is affected by educational factors, skills and competencies and job market factors. The literature review unveiled that the big data employees are lacking in the job market at a significant level due to various weaknesses in the aforementioned three factors. Similarly, the industry experts specifically revealed the lack of engagement with the profession locally. Hence, possible actions should be taken to uplift their representation as a prestigious career by mitigating all the shortcomings. The researchers determined that the local higher educational platforms and professional qualifications should be improved and standardized as the initial steps to mitigate the skill shortage of professionals. Similarly, the local employees should be extended with systematic competency development programmes in order to continuously nurture their skills. Meanwhile, the profession should be firmly embraced and promoted by local companies with the motive of a “win-win” approach to both the company and BDP. Their remuneration should be improved in line with the amount of value addition they bring to the company. Therefore, the companies should expand their horizons to grasp this enticing field as a source
of growth and competitive advantage.

5. REFERENCES


An investigation of factors affecting employability of big data professionals in Sri Lanka; with special reference to logistic companies


An investigation of factors affecting employability of big data professionals in Sri Lanka; with special reference to logistic companies


