The Similarities of Big Data Analytics on Company Performance to Track Procurement in Supply Chain Management

Authors

[1.] Dr. Amol Murgai, [2.] Dr. Bipllab Roy, [3.] Prof. Sumeet Bhatia, [4.] Dr. Purnendu Bikash Acharjee

[1.] Associate Professor, School of Business and Management,
Christ University, Lavasa, Pune
Email – amol.murgai@christuniversity.in
Mobile - 9890440576

[2.] Assistant Professor, School of Business and Management,
Christ University, Lavasa, Pune
Email – bipllab.roy@christuniversity.in
Mobile - 9101626663

[3.] Assistant Professor, School of Business and Management,
Christ University, Lavasa, Pune
Email – sumeet.bhatia@christuniversity.in
Mobile - 9810167567

[4.] Assistant Professor, School of Business and Management,
Christ University, Lavasa, Pune
Maharashtra State, India
Email – purnendu.acharjee@christuniversity.in
Mobile - 9864055678

Abstract: Big data analytics can add value and provide a new perspective by improving predictive analysis and modeling practices. This research is centered on supply-chain management and how big data analytics can help Yemeni supply-chain companies assess their experience, strategies, and professional capabilities in successfully implementing big data analytics, as well as assessing the tools needed to achieve these goals, including the results of implementation and performance achievement based on them. The research method used in the quantitative study was a sampling survey, using a questionnaire as a data collection tool. It included closed questions, measured with nominal and ordinal scales. A total of 205 managers provided complete and useful answers for this research. The collected data were analyzed with the Statistical Package for the Social Sciences (SPSS) package using frequency tables, contingency tables, and main component analysis. The major contributions of this research highlight the fact that companies are concerned with identifying new statistical methods, tools, and approaches, such as cloud computing and security technologies, that need to be rigorously explored.

Keywords: *supply-chain management; implementation; big data analytics; large industry*

1. Introduction

The provenance of structured and unstructured data makes it somewhat more difficult to analyze and generate information, but with supply-chain management systems, they aretransformed and tailored to end user requirements. The supply chain covers all activities from the development of production to production and logistics to maximize customer value and achieve sustainable competitive advantages.

Many companies still do not understand how to apply analytical techniques to achieve superior performance within the supply chain. Through this study, undertaken within Yemeni companies, we consider that we are making an important contribution to clarifying the aforementioned aspects. First, we identified the companies' experience in implementing big data analytics (BDA) in supply-chain management (SCM) and the difficulties encountered in this process, followed by the adoption of company strategies for implementing BDA in SCM.The two issues investigated are also related to the existence of professional capacities needed to develop information through BDA and to identify the supply-chain analysis tools used by companies and their future intention to implement new tools and technologies. Second, the paper article provides an in-depth understanding of the benefits and performance of companies after implementing BDA in SCM. Also, there have been some challenges, such as theacceptance and use of new technologies, as well as their regulation.

2. Literature Review

2.1. Big Data and Big Data Analytics

New challenges in capturing, collecting, analyzing, archiving, sharing, transferring, and processing large datasets in organizations led to the emergence of the BD concept. The determinant factor behind this concept is digitalization, with increased social and media popularity among electronic device users [3,4]. The big data concept has been applied to datasets that are very large and difficult to handle by traditional database management systems. In other words, their size exceeds the current capacity of software tools and storage systems for capturing, storing, managing, and processing data in an acceptable time [5].

According to specialists, the Big Data concept has different approaches, such as the “3Vs”: volume (the volume of current datasets in big data is a significant attribute, since such data is considered to be excluded from the traditional management techniques of databases); velocity(the rate at which data is collected); and variety (unstructured data are generated by sources such as social media, e-mails, and communication) [6,7].

Based on the 3V concept, other specialists define big data as follows: (1) “High-volume, high- speed, and/or large-scale computer equipment that requires cost-effective and innovative forms of information processing to enable improved understanding, decision-making and process automation” [8]; (2) “big data as volume, high speed and data of great variety used in decision-making and requiring innovative management techniques” [8]; and (3) “special type of large- scale data that cannot be stored, manipulated, and analyzed by means of a conventional system together with an anonymous source, various dimensions and its relationship cannot be easily measured due to its complexity and dynamic nature” [9].

The big data concept is expanded to “5Vs” by adding two more features (veracity, or reliability of collected data, and the value of datasets involving substantial investments). BDA is defined as a holistic approach for managing, processing, and analyzing data sizes (volume, variety, velocity, veracity, and value) that are needed to create action-oriented information for sustained delivery, performance measurement, and competitive advantage [10-12]. The 5Vs can be explained as follows: (1) volume refers to the magnitude of data that requires increased storage devices [13]; (2) variety is reflected by generating data from heterogeneous sources Internet of Things (IoT), online social networks, and structured, semi-structured, and unstructured formats) [14]; (3) velocity is given by the time to access, process, and use data in real time [11]; (4) veracity reflects the importance of data quality and confidence in the data used [15,16]; and (5) value is reflected by revealing unused data in big data and can support decision-making [17,18].

BDA involves the use of advanced analytical techniques for extracting important information from large volumes of data to facilitate decision-making [19]. Developed from the field of operational research, advanced analysis has had various classifications [20], among which descriptive, predictive, and prescriptive analysis are considered representative [21]. Descriptive analysis is based on the analysis of data describing past business situations, trends, patterns, and exceptions. The techniques used for descriptive analysis can be characterized as

[22] Standard reports and scoreboards, ad hoc reporting, query drilldown (OLAP) alerts, and viewing.

Predictive analysis is based on real-time data analysis and historical data to predict the likelihood of future events. This technology learns from existing data using machine learningtechniques and computational algorithms [23], including (1) advanced time series and advanced predictions used in SCM for marketing measures such as stockpiles or sales forecasts (ARIMA, ARMA); (2) supervised learning including linear and logistic regressions, statistical algorithms; (3) clustering; and (4) size reduction. Prescriptive analysis is based on data-basedpredictions to inform and suggest proposed action sets that can be advantages or avoid certain results and may include: (1) studies addressing the variability of expected outcomes by analyzing the scenario game theory; and (2) optimization and simulation under conditions of special relevance in the context of uncertainty based on computational stochastic programming of random variables (Monte Carlo).

2.2. Big Data Analytics in Supply-Chain Management

The SCM concept has been debated by specialists around the world. It has been defined as "management alongside and within a network of upstream and downstream organizations, bothof which have relationships and flows of material, information and resources" [24]. The supplychain can be considered to be a combination of four independent and interconnected entities (marketing, sourcing, inventory management, and transport). SCM is responsible for creating and maintaining links between the different entities in a business responsible for purchasing raw materials for final delivery of the product [25]. New technologies such as big data analytics synchronize SCM in a separate stream [26] and allow organizations to capture, process, analyze, store, and exchange data about their operations [27].

An extended supply chain is a multilayered system that connects organizations through collaboration and integration, as competition between supply chains is perceived as higher thanbetween individual firms [28]. Among the computer systems used for this purpose are Electronic Data Interchange (EDI), Vendor Managed Inventory (VMI), Efficient Consumer Response (ECR), Collaborative Planning Forecasting and Replenishment (CPFR), Collaborative Planning System (CPS), Sales Force Automation (SFA), Point of Sale (POS) data, and Customer Relationship Management (CRM) [29]. Among all SCM information flows, big data analytics focuses on data analysis and tools are included in the "analytics" domain. Analytics applies mathematics and statistics to large amounts of data. Big data without analytics is just a lot of data, and analytics without big data is simply math, statistical tools, and applications [30].

Thus, a first attempt to define BDA is based on the 3V framework: (1) weekly generated datavolume greater than 300 TB classifies companies in the BDA category; (2) the velocity of datacreation and transmission plays a key role in transferring data from batch processing to real- time operation; and (3) the variety of data provided to users can include, in addition to classical SQL or XML formats, digital information such as video, text, or images [7]. Starting from this attempt to define BDA within the supply chain, some specialists have presented other approaches: (1) the process of extracting and presenting supply-chain information to ensure measurement, monitoring, forecasting, and supply-chain management [31]; (2) a broad and unique view of the entire supply chain to disclose component or full production, including analyses and key performance indicators (KPIs) [32]; (3) using quantitative methods to obtainprospective information from data in order to understand in depth what is going upstream and downstream, in order to assess the operational decision-making impact of the supply chain [33]; (4) the operational leap from data response management models that can help specialists analyze larger datasets using proven analytical and mathematical techniques [34]; and (5) thecombination resulting from the application of quantitative and qualitative methods to SCM theory to solve SCM problems and predict results, taking into account the quality andavailability of data [35].

Studies on the impact of Big Data (BD) on sustainable investment in a supply chain (SC) have indicated the following aspects: (1) stimulation of the co-creation of value by reducing risks, with BD helping to shorten SC stages by extending economic marginalization and facilitating sustainable planning of smart investments in health care [36]; (2) a positive impact between the number of observations of market information and increased profit by using the updating information Bayesian approach of [37]; and (3) perfection of SC through investments made by both parties (suppliers and traders) in the profit generating BD and regulated by revenue distribution contracts between them [38].

An essential factor in motivating members within a SC to make sustainable investments in innovative technologies is related to equity concerns. They can promote and coordinate members of the SC, without the problem of advantageous inequity in view of the considerableinvestment in new sustainable innovative technologies [39]. A study was also conducted to explore the application of BDA in mitigating the social risk of a SC and how it can contribute to achieving ecological, economic, and social sustainability. The results indicated that companies can predict various social problems (labor safety, fuel consumption monitoring, health of the workforce, security, physical condition of vehicles, ethical behavior, theft, speed and traffic violations through big data analysis) that can be managed through the information provided, thus contributing to the mitigation of social risks [40].

The influence of block-chain technology on the exchange of information between participants in a SC is highlighted with the help of mean-Conditional Value at Risk (CVaR) to characterize the risk-aversion behavior of the traders. The CVaR-based revenue distribution contract is used to coordinate the SC and profit distribution. Research carried out by specialists indicated that profits from the SC system are higher in a centralized than a decentralized decision-making process. Through block-chain technology, transaction costs between SC members can be reduced, information exchange can be achieved, and benefits can be improved [41].

Using game theory, another study dedicated to investigating the risk-aversion behavior of producers and traders within a closed-loop SC examined making optimal decisions about wholesale, retail, and recycling prices in centralized and decentralized decision scenarios. By analyzing some parameters that influence the revenue distribution contract, a new contract model was proposed to increase the SC members' profits by identifying the optimal income distribution ratio [42].

Other equally interesting studies were devoted to the research and development of a new model based on the Decision-Making Trial and Evaluation Laboratory (DEMATEL) method called the KTT-GSCM (Knowledge-Technology Transfer-Green Supply-Chain Management) multicriteria model which demonstrated the mutual influences between GSCM, KTT, and technological innovation [43]. Based on a relational vision of inter-organizational competitive advantage, another study focused on empirical research of drivers and their results in integrating a low-carbon supply chain (logistics service providers (LSPs) with supply-chain integration (SCI)). The results of the study highlighted the fact that between the corporate environmental responsibility of the LSP and the environmental requirements of the clients, there are positive relationships with SCI with the low-carbon emissions of the LSP, and the latter is positively related to the financial and environmental performance of the LSP [44].

To ensure environmentally sustainable logistics, companies must have an environmentally sustainable logistics performance management (ESLPM) process. Transposing the integrationof processes within the SC to increase performance was achieved by developing a frameworkaimed at integrating the ESPLM process and third-party logistics (3PL). This framework can provide guidance for practitioners in identifying the degree of integration of logistics performance management processes and decision-making at the senior level [45].

Large companies (with more than 250 employees) are already using supply-chain analytics tools such as KPMG Spectrum Third-Party Intelligence, Deloitte Supply-Chain Solutions, Halo Supply-Chain Analytics and Business Intelligence Software, Supply-Chain and Logistics Analytics Software. Numerous software solutions are also in use, such as Relax, FusionOps, Blue Ridge Supply-Chain Analytics and Inventory Analytics Software, IMI Supply-Chain Analytics, Supply-Chain Analytics [29]. These companies manage large volumes of data of thousands of Small and medium-sized enterprises (SMEs).

2.3. Importance and Application of Big Data in Supply Chain Management

The technological advances recorded by supply-chain entities, the volume of data generated, and the speed of distribution have led to significant increases in structured and unstructured data analysis to get a clearer picture of customer needs and improve cost-related aspects of supply-chain processes. Big data analytics can make significant contributions to areas such as product development, market demand prediction, supply decisions, distribution optimization,

and customer feedback. Companies with a disciplined strategy of using big data analytics have had better results with investments [46]. In other words, a clear and systematic strategy of big data analytics can provide a better return on investment (ROI) in certain areas of the supply chain, such as marketing, purchasing, shipping, and storage [47].

Other specialists have demonstrated the importance and contribution of BDA to SCM by: (1) improving manufacturing performance by linking IoT and BD to manufacturing systems to minimize bottlenecks by developing forecasting techniques [48]; (2) observing current trends in supply-chain management by using Twitter and developing a new conceptual framework inthis regard [49]; (3) investigating the potential use of TD for the management of product life cycles [50]; (4) measuring the sustainability of supply chains using BD prediction analysis [15]; and (5) establishing a relationship between sustainable supply-chain management and logistics operations in the food industry [51], and developing a methodology for analyzing social data for supply-chain analysis and logistics operations in the food industry [52].

From the above approaches we can see that: (1) the information from SCM analysis should bepresented and extracted in a way that will help end users; (2) enhanced data integration and SCM analysis have helped to increase visibility across the supply chain; (3) data retrieval methods and their permanent updates have helped to improve the speed of data processing, with real-time capability for various decision situations; (4) data analysis has been forward- looking and the impact assessment of prospective decisions has led to the emergence of a new advanced supply-chain management using proactive models; and (5) there is a need to includeknowledge in the use and analysis of data. In other words, BDA is the natural evolution of bigdata in SCM [53].

2.4. Benefits and Constraints of BDA for SCM

Due to the upsurge of unstructured data in complex processes across entities, big data analyticshas become the biggest challenge for the supply chain. The competitive position on the market is maintained under the conditions of proper management of the entities in the supply chain. Entities are interconnected through material flows, financial flows, and significant electronic information exchanged simultaneously among all supply-chain partners. Connections between different parts or elements of the supply chain may be direct or indirect, and significant interactions between them determine the complexity of the system.

To understand this impact, we need to take into account the following key features: (1) the number of entities in the supply chain; (2) the diversity of entities; (3) the existing interdependence between items, products, and supply-chain partners; (4) the dynamics of the system, represented by variety; and (5) the existence of uncertainties. The great interest of specialists is shown by the empirical studies undertaken, which highlight several advantages of BDA within EMS, such as reducing operational costs, increasing customer satisfaction, andincreasing SC agility [35,54-56].

In previous studies of specialists, numerous benefits related to SCM information technology (IT) optimization have been identified, such as (1) the exchange of information between SC stakeholders [57-60]; (2) the transformation of intra- and inter-organizational business processes (cancellation, redesign, automation) [61-63], operational efficiency [64] and revenuegrowth [59], profitability, and improvement of stakeholder relations [65,66]; and

(3) improving relational and contractual governance by effectively mitigating the opportunism of partners [67]. In addition, BDA could also generate future opportunities for stakeholders such as obtaining a competitive advantage [68] and reducing exposure to fraud or other malfunctions [69,70].

According to specialist studies, several constraints have been identified in the adoption/implementation of BDA in SCM: technological barriers, expertise, and investment barriers, data barriers, organizational barriers, etc. In the category of technological barriers, the following issues were identified as the basis for BDA implementation in SCM: (1) a lack of understanding of the implementation of new technologies or a lack of tools needed to implement BDA in SCM [54]; (2) a lack of infrastructure facilities [71-73]; (3) low acceptance,routine, and BDA assimilation by SCM organizations and partners [73,74]; and (4) data security [35,75].

With regard to expertise and investment barriers, specialist studies have shown interesting aspects of BDA in SCM: (1) a lack of qualified IT staff and high investment cost [72,73]; and

(2) A lack of funds and facilities for research and development of BDA instruments [70].

Databarriers have highlighted interesting aspects of the implementation of BDA in SCM, such as:

(a) Complexity of data integration and data quality [72,73,76]; and (2) data security, confidentiality, performance, and scalability [72,73]. Organizational barriers have been identified as negative limits in implementing BDA in SCM: (1) the absence of a data-sharing policy between organizations and a lack of thinking in terms of large data [69]; and (b) a lackof training facilities and time constraints [72,77].

2.5. Big Data Analytics and Supply-Chain Management in Yemen

Studies in Yemen related to the implementation of BD and BDA are scarce. The most significant studies were in the fields of health and information and communication technology. BD can help physicians choose treatments more correctly and faster, based on information collected by health care professionals. Thus, patients can benefit from appropriate treatment indue time, and will be better informed about health care providers [78]. In the field of information and communication technology, the themes of large data management and management and analysis (big data) have been analyzed in terms of their relevance for the solutions they offer to increase competitiveness in intelligent specialization at a national level.

On the other hand, the role of Research, Development and Innovation (RDI) in information and communication technology (ICT) in supporting the development of the economy and society, focusing on the business environment, identifies among the four priorities e-commerce, research and development (R&D) and innovation in ICT.

In line with these priorities, in the Digital Agenda for Yemen Program, the electronic services section includes the project "ICT Research and Development and Innovation: Developing innovative products and services serving the 10 sectors (tourism and ecotourism, wood and furniture, creative industries, automotive and components, information and communication technology, food and beverage processing, health and pharmaceuticals, energy and environmental management, bio-economy, bio pharmaceutics, and biotechnology) identified in smart specialization" (TIC-SI), with the objective of investigating and implementing this role.

In Yemen, the SC concept is widely used, having being given more meanings, such as:

(1) the variation and nonexistence of a uniform designation between SC and SCM; and

(2) the use of synonyms for SC (supply network, logistics network, supply-chain management, supply-chain provision). Supply-chain management involves challenges such as building trust, sharing information on market needs, developing new products and services, and meeting customer requirements as efficiently as possible.

The starting point of research regarding the stage of implementation, expansion, and development of SCM in Yemeni companies is the logistics sector. Despite with a background of poor transport infrastructure and public policy and an economic crisis, this sector managed to make some progress, justified by the presence of large companies and international groups in Yemen. This can also be noticed by tracking foreign investments in the economy following the large central areas (Bucharest, Cluj-Napoca, Arad, Constanta, and Ploiesti) compared to the performance level of logistics platforms of large companies in developed countries such as the US or European countries. Yemeni logistics activities are carried out internally with specialized companies, as well as externally (the tendency is increasing, with outsourcing to specialized suppliers). During 2008, the stage of development of SCM was low, with only extension and integration of companies with suppliers and distributors.

The implementation of SCM within Yemeni companies is influenced by two major factors:

(1) Physical capital (technology not updated) and poor human capital (reduction of wage

costs); and limited vision of the inter-organizational structure. The directions for improvement have been investment in new technology and human capital, assimilation of new values by managers and continuous adaptation, and improvement of employees' skills and their adaptation to new technologies.

The implementation of SCM within Yemeni companies requires organizational changes based on the "Eight I's that Create We's" approach, which considers the following characteristics: individual excellence, importance, interdependence, investment, information, integration, institutionalization, and integrity. Unlike later research in Yemen, our paper aims to explore aspects related to the state of implementation of BDA for SCM, the adoption of BDA strategies in SCM, the identification of BDA's capabilities and tools in SCM, and the measurement of future intentions of Yemeni companies regarding the implementation of otherBDA tools for SCM.

3. Research Methodology

Given the importance of using Industry 4.0's new tools and technologies, which make a substantial contribution to improving business performance, we conducted a study of companies of Yemen. Industry 4.0 represents the fourth industrial revolution with a major impact on the production of the future, which integrates innovative elements and technologies such as big data analytics, Internet of things and others, and which ensures constant communication and connection in addressing customer services. Through this study, we wanted to find out the future intention of companies to implement new tools and technologies that impact their performance.

The quantitative market research had the following objectives:

1: Identify the companies' experience in implementing analytics in supply chainand the difficulties encountered in this process.

2: Identify the supply-chain analytics tools used by companies and the futureintended implementation of new tools and technologies to gain valuable supply-chain insights.

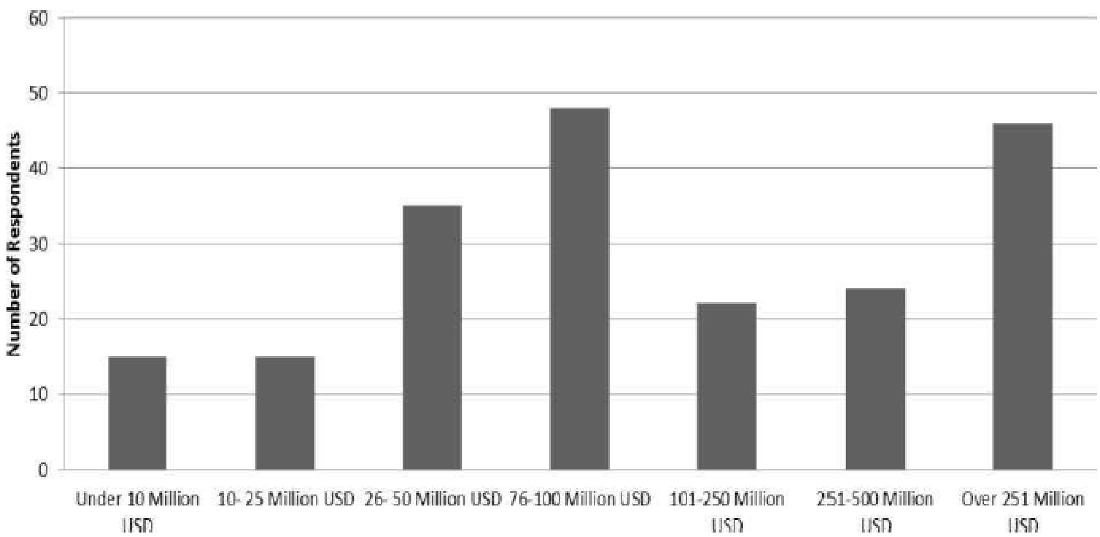
The following research hypothesis was formulated:

Ho: There is no link between the size of the company and its experience in implementing analytics in supply chain.

To achieve the above-mentioned objectives, but also to test the model proposal, it was necessary to collect data from different companies. The study was conducted between

January and April 2022. The relevant population for our study was identified in the metadata database of the National Institute of Statistics in Yemen. Starting from the information provided, the sampling base was defined and built. The process ended with a cross-listing of 40 companies. In view of some errors (such as inactive, noncontact, or already dissolved), the sampling base(list) was reduced to 25 enterprises. After companies were identified, contact with their managers was established to receive the survey agreement and know which managers were toget the online questionnaire, included in an e-mail link to be completed later.

The research method used in the quantitative study was a survey by sampling, using the questionnaire as the data collection tool. It included closed questions, measured with nominaland ordinal scales. The study was carried out with the support of interviewers with experience in the field. Managers who agreed to participate in the survey received an electronic link via e-mail for the online questionnaire. The process ended with the conclusion ofparticipation agreements with 20 companies included in the list. The response rate was 90.7%,i.e., 20 managers provided complete and useful answers for this research. The collected data were analyzed with the SPSS package, using the frequency and contingency tables, chi- square test, the Student's t-test, factorial correspondence analysis, and the binary logistic regression model.



Annual Revenue in USD Figure 5.1: Annual Revenue in USD and Number of Respondents

Chart 01 – Companies with annual revenue considered for research (Source: Primary Data)

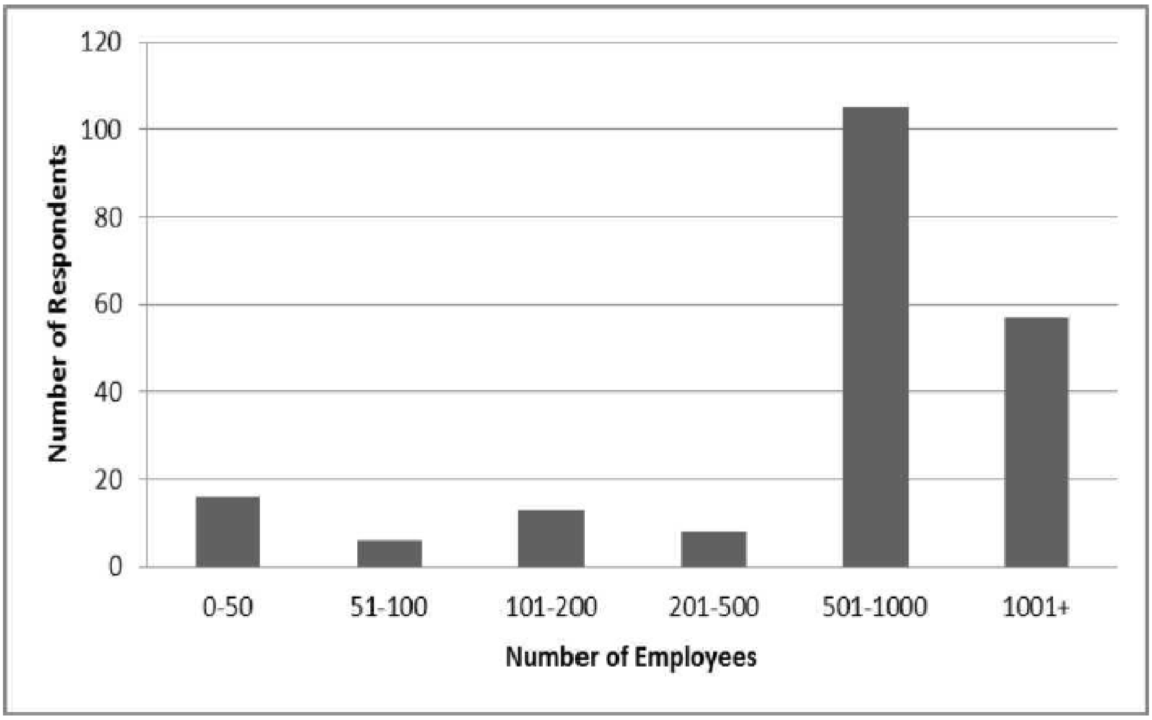
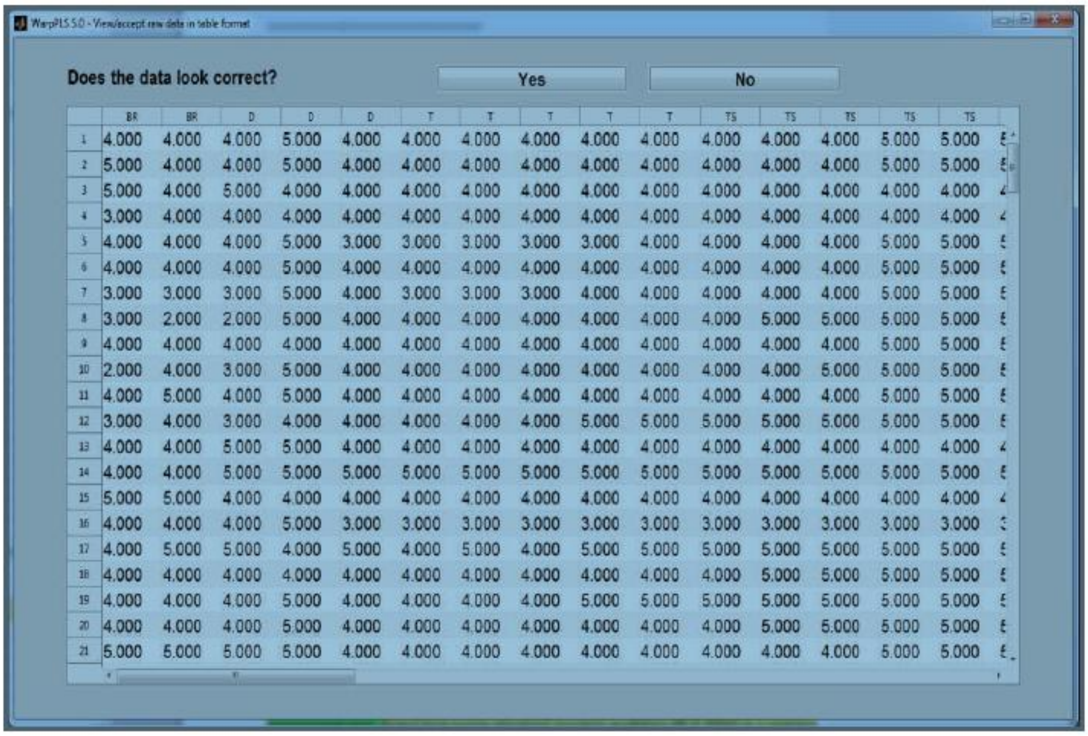


Figure 5.2: Employees Strength and Number of Respondents in Surveyed Firms

Chart 02 – Companies with employee strength considered for research (Source: Primary Data)



	BR	BR	D	D	D	T	T	T	T	T	TS	TS	TS	TS	TS
1	4.000	4.000	4.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000
2	5.000	4.000	4.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000
3	5.000	4.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000
4	3.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000
5	4.000	4.000	4.000	5.000	3.000	3.000	3.000	3.000	3.000	4.000	4.000	4.000	4.000	5.000	5.000
6	4.000	4.000	4.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000
7	3.000	3.000	3.000	5.000	4.000	3.000	3.000	3.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000
8	3.000	2.000	2.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000	5.000	5.000
9	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000
10	2.000	4.000	3.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000	5.000
11	4.000	5.000	4.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000
12	3.000	4.000	3.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000
13	4.000	4.000	5.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000
14	4.000	4.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000
15	5.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000
16	4.000	4.000	4.000	5.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000
17	4.000	5.000	5.000	4.000	5.000	4.000	5.000	4.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000
18	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000	5.000	5.000
19	4.000	4.000	4.000	5.000	4.000	4.000	4.000	4.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000
20	4.000	4.000	4.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000	5.000	5.000
21	5.000	5.000	5.000	5.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	5.000	5.000

Chart 03 – WarpPLS SEM Screen for Raw Data Review and Acceptance (Source: Primary Data)

4. Results and Discussions

The first research goal was to identify the companies' experience in implementing BDA in SCM and the difficulties encountered in this process. Using the chi-square test, we can see the distribution of companies according to the experience gained in implementing BDA in the SC. In total, there are more than four times as many experienced companies as companies with no experience.

From the analysis of the differences between observed and expected frequencies, it can be seen that there are differences in the levels of all subgroups formed by the crossing of two variables. In the subgroups of companies with more than 100 employees, observed experience is higher than expected and in companies with lack of experience in implementing analytics in the supply chain, expected frequency is higher than observed

5. Conclusions

Massive volumes of data coming from different sources have a positive effect on real-time decision-making. The variety of data sources, the quality of the data to be integrated, and their visualization are some of the challenges for big data analytics integration. The survey results indicate that 80% of Yemeni companies have accumulated big data analytics experience, integrating different software solutions into the supply chain. Big companies (with more than 100 employees) allocate large annual budgets to undertake projects aimed at implementing big data analytics in the supply chain or to employ specialists in the field. The difficulties encountered in the implementation of big data analytics in the supply chain by companies with annual sales revenues of up to €10 million are related to large investment requirements, security issues, and lack of executive support. Entities with annual sales over €11 million have encountered obstacles related to security and privacy issues.

Moreover, 90% of respondents have adopted an enterprise-wide strategy (that includes the supply chain), facilitating the use of big data to add business value. Some of these entities have adopted a dedicated supply-chain strategy (35.6%), and 11.2% have applied a big data strategy in some form for some processes. Some manufacturing (24.2%) and e-commerce (4.4%) companies have adopted a strategy that focuses on improving predictions of customer needs, while consulting companies (4.4%) have as a priority improving the efficiency of the supply chain. To develop insights through big data analytics, Yemeni companies need professional capabilities; about 91.22% of entities use such capacities and are 5-10 years old. Of these, 43.9% have a specialized team using sophisticated tools to generate insights, 31.2% have an independent team, and 16.1% have a traditional database (part of the IT team).

Company managers admit that technological changes have changed work processes lately, with investments in training and staff development for large data analytics needed. Companies with over five years of experience in the market that have benefited from the expertise of specialists have achieved results including “Improvement in customer service and demand fulfillment,” “Improved cost to serve,” “Increase in supply-chain efficiency,” and “Optimization of inventory and asset productivity.” Other companies that are smaller and newer in the Yemeni landscape, benefiting from the support of teams of specialists, have obtained results related to “Greater integration across the supply chain,” “Faster and more efficient response time to supply-chain issues,” and “Optimization of inventory and asset productivity.”

Managers agree that new capabilities and technologies are needed to transform, manage, and analyze company-wide information. The main challenges they face are the acceptance and use of new technologies, as well as their regulation. The most notable problems to overcome are based on the difficulty of analyzing large volumes of data to achieve timely accurate results, and the need for standardization, interoperability, security, confidentiality, expertise, and funding to develop big data analytics infrastructure and integrate sets of already available data.

Managers recognize that their concerns include identifying new methods, tools, and statisticaltechnology approaches, such as cloud computing and security technologies, to be rigorously explored. Big data analytics is an opportunity to use new types of data to create more agile businesses to solve problems that were previously considered unsolvable, leading to better business results. This will lead to radical changes in business operations that change from theuse of a model based mainly on the experience of decision-makers to an information model that gives real value to the business and organization itself.

Starting from the objectives presented and analyzed in our study, we propose the following possible future research directions: (1) Studying the opportunities, challenges, advantages and disadvantages of BD in large companies and/or SMEs in the state or private sector; (2)

Studying how BD systematically affects the economic value in the business environment; (3) Studying the implementation of BD in various sectors of activity and the efficiency of using the information obtained in making decision process; (4) Studying the capabilities and benefits of using BDA in optimizing SCM; (5) Studying the impact of BDA on SCM by using varioustools dedicated to analyzing information.

References

1. Turkes, M.C.; Oncioiu, I.; Aslam, H.D.; Marin-Pantelescu, A.; Topor, D.I.; Capusneanu,
S. Drivers and barriers in using industry 4.0: A perspective of SMEs in Yemen. Processes **2019**, 7, 153.
2. Manyika, J.; Chui, M.; Brown, B.; Bughin, J.; Dobbs, R.; Roxburgh, C.; Byers, A.H. BigData: The Next Frontier for Innovation, Competition and Productivity; Mckinsey GlobalInstitute: New York, USA, 2016.
3. Hellerstein, J. Parallel Programming in the Age of Big Data.
4. Lohr, S. The Age of Big Data. The New York Times. 2012.
5. Kubick, W.R. Big data, information and meaning. Appl. Clin. Trials **2012**, 21, 26-28.
6. Wisner, J.; Tan, K.C.; Leong, G.K. Principles of Supply Chain Management: A BalancedApproach, 4th ed.; Cengage Learning: Boston, MA, USA, 2012.
7. Laney, D. 3D data management: Controlling data volume, velocity and variety. META Group Res. Note **2001**, 6, 1.
8. Beyer, M.A.; Laney, D. The Importance of "Big Data": A Definition; Gartner Publications: Stamford, CT, USA, 2012.
9. Sun, E.W.; Chen, Y.T.; Yu, M.T. Generalized optimal wavelet decomposing algorithm forbig financial data. Int. J. Prod. Econ. **2015**, 165, 194-214.
10. Wamba, S.F.; Akter, S.; Edwards, A.; Chopin, G.; Gnanzou, D. How 'big data' can makebig impact: Findings from a systematic review and a longitudinal case study. Int. J. Prod.Econ. **2015**, 165, 234-246.
11. Assungao, M.D.; Calheiros, R.N.; Bianchi, S.; Netto, M.A.S.; Buyya, R. Big data computing and clouds: Trends and future directions. J. Parallel Distrib. Comput. **2015**, 79,3-15.
12. Emani, C.K.; Cullot, N.; Nicolle, C. Understandable big data: A survey. Comput. Sci. Rev. **2015** , 17 , 70-81.

13. Chen, C.L.P.; Zhang, C.Y. Data-intensive applications, challenges, techniques and technologies: A survey on big data. Inf. Sci. **2014**, 275, 314-347.
14. Tan, K.H.; Zhan, Y.; Ji, G.; Ye, F.; Chang, C. Harvesting big data to enhance supply chaininnovation capabilities: An analytic infrastructure based on deduction graph. Int. J. Prod.Econ. **2015**, 165, 223-233.
15. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. Int.J. Inf. Manag. **2015**, 35, 137-144.
16. White, M. Digital workplaces: Vision and reality. Bus. Inf. Rev. **2012** , 29 , 205-214
17. Lee, A.H.I.; Kang, H.-Y.; Ye, S.-J.; Wu, W.-Y. An integrated approach for sustainable supply chain management with replenishment, transportation, and production decisions. Sustainability **2018**, 10, 3887.
18. Oracle. Big Data for the Enterprise ; Oracle: Redvud, CA, USA, 2012.
19. Tsai, C.-W.; Lai, C.-F.; Chao, H.-C.; Vasilakos, A.V. Big data analytics: A survey. J. BigData **2015**, 2, 21.
20. Chae, B.; Sheu, C.; Yang, C.; Olson, D. The impact of advanced analytics and data accuracy on operational performance: A contingent resource based theory (RBT) perspective. Decis. Support Syst. **2014**, 59, 119-126.
21. Lustig, I.; Dietrich, B.; Johnson, C.; Dziekan, C. The analytics journey. Anal. Mag. **2010**,3,11-13.
22. Zeng, X.; Lin, D.; Xu, Q. Query performance tuning in supply chain analytics. In Proceedings of the 4th International Conference on Computational Sciences and Optimization (CSO), Kunming and Lijang City, China,15-19April2011;p.327.
23. Siegel, E. Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die; Wiley Publishing: Hoboken, NJ, USA, 2013.
24. Christopher, M. Logistics & Supply Chain Management, 4th ed.; FT Prentice Hall: UpperSaddle River, NY, USA, 2011.
25. Halo. Descriptive, Predictive, and Prescriptive Analytics Explained. 2018.
26. Edwards, P.; Peters, M.; Sharman, G. The effectiveness of information systems in supporting the extended supply chain. J. Bus. Logist. **2001**, 22, 1-27.
27. Smith, G.E.; Watson, K.J.; Baker, W.H.; Pokorski, J.A. A critical balance: Collaborationand security in the IT-enabled supply chain. Int. J. Prod. Res. **2007**, 45, 2595-2613.
28. Antai, I.; Olson, H. Interaction: A new focus for supply chain vs. supply chain competition. Int. J. Phys. Distrib. Logist. Manag. **2013**, 43, 511-528.

29. Barratt, M.; Oke, A. Antecedents of supply chain visibility in retail supply chains: A resource-based theory perspective. J. Oper. Manag. **2007**, 25, 1217-1233.
30. Sanders, N.R. Big Data Driven Supply Chain Management: A Framework for Implementing Analytics and Turning Information into Intelligence, 1st ed.;Pearson: Hoboken, NJ, USA, 2014.
31. Marabotti, D. Build supplier metrics, build better product. Quality **2003**, 42, 40-43.
32. Sahay, B.S.; Ranjan, J. Real time business intelligence in supply chain analytics. Inf. Manag. Comput. **2008**, 16, 28-48.
33. Niu, S.; Zhuo, H.; Xue, K. DfRem-driven closed-loop supply chain decision-making: A systematic framework for modeling research. Sustainability **2019**, 11, 3299.
34. O'Dwyer, J.; Renner, R. The promise of advanced supply chain analytics. Supply Chain Manag. Rev. **2011**, 15, 32-37.
35. Moro Visconti, R.; Morea, D. Big data for the sustainability of healthcare project financing. Sustainability **2019**, 11, 3748.
36. Cheng, Y.; Kuang, Y.; Shi, X.; Dong, C. Sustainable investment in a supply chain in the big data era: An information updating approach. Sustainability **2018**, 10, 403.
37. Xu, L.; Gao, R.; Xie, Y.; Du, P. To be or not to be? Big data business investment decision-making in the supply chain. Sustainability **2019**, 11, 2298.
38. Du, B.; Liu, Q.; Li, G. Coordinating leader-follower supply chain with sustainable green technology innovation on their fairness concerns. Int. J. Environ. Res. Public Health **2017**,14, 1357.
39. Mani, V.; Delgado, C.; Hazen, B.T.; Patel, P. Mitigating supply chain risk via sustainability using big data analytics: Evidence from the manufacturing supply chain. Sustainability **2017**, 9, 608.
40. Liu, L.; Li, F.; Qi, E. Research on risk avoidance and coordination of supply chain subjectbased on blockchain technology. Sustainability **2019**, 11, 2182.
41. Zou, H.; Qin, J.; Yang, P.; Dai, B. A coordinated revenue-sharing model for a sustainableclosed-loop supply chain. Sustainability **2018**, 10, 3198.
42. Pinto, M.M.A.; Kovaleski, J.L.; Yoshino, R.T.; Pagani, R.N. Knowledge and technology transfer influencing the process of innovation in green supply chain management: A multicriteria model based on the DEMATEL Method. Sustainability **2019**, 11, 3485.
43. Qian, C.; Wang, S.; Liu, X.; Zhang, X. Low-carbon initiatives of logistics service

providers: The perspective of supply chain integration. Sustainability **2019**, 11, 3233.

44. Persdotter Isaksson, M.; Hulthen, H.; Forslund, H. Environmentally sustainable logistics performance management process integration between Buyers and 3PLs. Sustainability **2019**, 11, 3061.
45. Waller, M.A.; Fawcett, S.E. Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. J. Bus. Logist. **2013**, 34, 77-84.
46. Accenture. Accenture Big Success with Big Data Survey. 2014. Available online:https://www.slideshare.net/polenumerique33/accenture-bigdatapov1 (accessed on20 June 2019).
47. Benabdellah, A.C.; Benghabrit, A.; Bouhaddou, I.; Zemmouri, E.M. Big data for supply chain management: Opportunities and challenges. In Proceedings of the IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), Agadir, Morocco, 29 November—2 December 2016.
48. Bi, Z.; Cochran, D. Big data analytics with applications. J. Manag. Anal. **2014**, 1, 249 - 265.
49. Chae, B. Insights from hashtag #supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research. Int. J. Prod. Econ. **2015**, 165, 247-259.
50. Li, J.; Tao, F.; Cheng, Y.; Zhao, L. Big data in product lifecycle management. Int. J. Adv.Manuf. Technol. **2015**, 81, 667-684.
51. Hazen, B.T.; Boone, C.A.; Ezell, J.D.; Jones-Farmer, L.A. Data quality for data science, predictive analytics, big data in supply chain management: An introduction to the problem & suggestions for research and applications. Int. J. Prod. Econ. **2014**, 154, 72-80.
52. Singh, A.; Shukla, N.; Mishra, N. Social media data analytics to improve supply chain management in food industries. Transp. Res. Part E Logist. Transp. Rev. **2018**, 114, 398 - 415.
53. Rozados, I.V.; Tjahjono, B. Big data analytics in supply chain management: Trends and related research. In Proceedings of the 6th International Conference on Operations and Supply Chain Management, Bali, Indonesia, 10-12 December 2014.
54. Schoenherr, T.; Speier-Pero, C. Data science, predictive analytics, and big data in supplychain management: Current state and future potential. J. Bus. Logist. **2015**, 36,

120-132.

55. Sheffi, Y. Preparing for disruptions through early detection preparing for disruptions through early detection. MIT Sloan Manag. Rev. **2015**, 57, 31-42.
56. Ramanathan, U.; Subramanian, N.; Parrott, G. Role of social media in retail network operations and marketing to enhance customer satisfaction. Int. J. Oper. Prod. Manag. **2017**, 37, 105-123.
57. Sahin, F.; Robinson, E.P. Flow coordination & information sharing in supply chains: Review, implications, directions for future research. **2002**, 33, 505-536.
58. Saeed, K.A.; Malhotra, M.K.; Grover, V. Examining the impact of inter organizational systems on process Efficiency and sourcing leverage in buyer-supplier dyads. Decis. Sci.**2005**, 36, 365-396.
59. Rai, A.; Patnayakuni, R.; Seth, N. Firm performance impacts of digitally enabled supply chain integration capabilities. MIS Q. **2006**, 30, 225-246.
60. Eric, T.G.W.; Hsiao-Lan, W. Inter organizational governance value creation: Coordinating for information visibility and flexibility in supply chains. Decis. Sci. **2007**,38, 647-674.
61. Wamba, F.S.; Lefebvre, L.A.; Bendavid, Y.; Lefebvre, E. Exploring the impact of RFID technology and the EPC network on mobile B2B eCommerce: A case study in the retail industry. Int. J. Prod. Econ. **2008**, 112, 614-629. [CrossRef]
62. Wamba, F.S. Achieving supply chain integration using RFID technology: The case of emerging intelligent B-to-B e-commerce processes in a living laboratory. Bus. Process Manag. J. **2012**, 18, 58-81.
63. Asoo, J.V. E-business and supply chain management. Decis. Sci. **2002**, 33, 495-504.
64. Devaraj, S.; Krajewski, L.; Wei, J.C. Impact of e-Business technologies on operational performance: The role of production information integration in the supply chain. J. Oper.Manag. **2007**, 25, 1199-1216.
65. Sethuraman, M.S. Big Data's Impact on the Data Supply Chain. In *Cognizant 20-20 Insights;* Cognizant: Teaneck, NJ, USA, 2012.
66. Zhang, T. How do information technology resources facilitate relational and contractual governance in green supply chain management? *Sustainability* **2019**, *11*, 3663.
67. Vasan, S. Impact of Big Data and Analytics in Supply Chain Execution. Supply Chain Digital. 2014.

68. Hagstrom, M. High-performance analytics fuels innovation and inclusive growth: Use big data, hyper connectivity and speed to intelligence to get true value in the digital economy.*J. Adv. Anal.* **2012**, *2*, 3-4.
69. Kenny, J. Big Data Can Have Big Impact on Supply Chain Management: The Use of Data Analytics Is Underused in Supply Chain Management to Minimize Risk Exposure; IC Inside Counsel. 2014.
70. Moktadir, M.A.; Mithun Ali, S.; Kumar, P.S.; Shukla, N. Barriers to big data analytics in manufacturing supply chains: A case study from Bangladesh. *Comput. Ind. Eng.* **2019**, *128*, 1063-1075.
71. Trelles, O.; Prins, P.; Snir, M.; Jansen, R.C. Big data, but are we ready? *Nat. Rev. Genet.* **2011**, *12*, 224.
72. Malaka, I.; Brown, I. Challenges to the organisational adoption of big data analytics: A case study in the South African telecommunications industry. In Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists, Stellenbosch, South Africa, 28-30 September 2015; pp. 1-9.
73. Alharthi, A.; Krotov, V.; Bowman, M. Addressing barriers to big data. *Bus. Horiz.* **2017**, *60*, 285-292.
74. Gunasekaran, A.; Papadopoulos, T.; Dubey, R.; Wamba, S.F.; Childe, S.J.; Hazen, B.; Akter, S. Big data and predictive analytics for supply chain and organizational performance. *J. Bus. Res.* **2017**, *70*, 308-317.
75. Dubey, R.; Gunasekaran, A.; Childe, S.J.; Wamba, S.F.; Papadopoulos, T. The impact of big data on world-class sustainable manufacturing. *Int. J. Adv. Manuf. Technol.* **2016**, *84*,631-645.
76. Fallik, D. For big data, big questions remain. *Health A*ff. **2014**, *33*, 1111-1114.
77. Zhong, R.Y.; Newman, S.T.; Huang, G.Q.; Lan, S. Big data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Comput. Ind. Eng.* **2016**, *101*, 572-591.
78. Alexandru, A.; Coardos, D. Utilizarea Tehnologiilor Big Data si IoT in Domeniul Sanatatii. *Rev. Romana Inform. Autom.* **2018**, *28*, 61-84.