IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 7.301

Volume 3, Issue 3, January 2023

Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management

Ms. Dalvi Swapnali[1] and Mrs. Akshata Chavan[2]

Student, M.Sc. I.T., I. C. S. College, Khed, Ratnagiri, Maharashtra, India[1]

Asst. Prof., Department of I.T., I. C. S. College, Khed, Ratnagiri, Maharashtra, India[2]

Abstract: *The supply chain management (SCM), data science, predictive analytics, and big data (together referred to as DPB) confluence offers a wealth of study options. We show how the increased use of these terminology might help supply chain education and research. Data science necessitates both domain expertise and a broad range of quantitative skills, despite the paucity of research on the topic and the abundance of open issues. We suggest further study into the competencies required of SCM data scientists and examine the relationship between domain expertise and the efficacy of SCM data scientists. This expertise is crucial for the advancement of future supply chain executives. We suggest data science and predictive analytics definitions that are particular to SCM. We examine real-world instances of DPB uses and propose both DPB-based research issues derived from these applications and from management theories. Last but not least, we provide a detailed explanation of the steps researchers might take to respond to our request for research on the confluence of SCM and DPB.*

Keywords: Data Science; Predictive Analytics; Big Data; Logistics; Supply Chain Management; Design; Collaboration; Integration; Education

I. INTRODUCTION

The current buzzword is "big data."Big data, on the other hand, has the potential to alter the design of business models as well as the day-to-day decision-making that goes along with emerging data analysis. This is more than the typical faddish fuzz. Supply chain management (SCM) faces both a huge opportunity and a significant challenge as a result of this expanding mix of resources, tools, and applications. According to Mayer-Schonberger and Cukier (2013), big data are being used to transform medical practice, modernize public policy, and guide business decision-making. In fact, more data have been recorded in the last two years than in all of human history. Supply chain dynamics have the potential to be transformed by big data.

Data sets that are larger than can be managed by traditional, hands-on management tools are the result of an increase in the quantity and variety of data. New predictive analytics applications and data science methods have been developed to manage these new, potentially valuable data sets. This brand-new intersection of big data, predictive analytics, and data science will be referred to as DPB.

There is some evidence to support the widespread belief that data contribute to improved decision-making and profitability. According to McAfee and Brynjolfsson (2012), who conducted a large-scale study, "the more companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational results... companies in the top third of their industry in the use of data-driven decision making were on average, 5% more productive and 6% more profit- able than their competitors" Supply chain researchers and managers must comprehend and embrace DPB's role and implications for decision-making in the supply chain in order to benefit from the big-data revolution.

II. DATA SCIENCE, PREDICTIVE ANALYTICS AND BIGDATA

DPB is getting more attention from the public, business, and academia. For instance, the Harvard Business Review published three articles in its October 2012 issue that are relevant to this editorial: Large Data: Data Scientist: The Management Revolution" (McAfee and Brynjolfsson, 2012),"Making Advanced Analytics Work for You" by Barton

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 7.301
Volume 3, Issue 3, January 2023

and Court (2012) and "The Sexiest Job of the 21st Century" by Davenport and Patil (2012).The lead article in MIS Quarterly's special issue on business intelligence was titled "Business Intelligence and Analytics:From Big Data to Big Change" (Chen et al.2012).On these subjects, there are also a lot of articles in trade and even general publications.
Even a brand-new journal, Big Data, which debuted in March 2013, exists.
We have attempted to comprehend the DPB's implications for business logistics and SCM education and research over the past few years. We believe that these new tools will fundamentally alter the design and management of supply chains, posing a novel and significant challenge to SCM and logistics. To meet this challenge, research and education priorities may need to shift. Rethinking a lot of conventional approaches will be necessary. In the new data-rich environment, some standard procedures might even be discarded as being out of date. Some individuals might view the possibilities as threats rather than as opportunities. DPB, on the other hand, is fundamentally compatible with SCM, so the immense value of DPB is within our grasp.

SCM data scientist skill set

Discipline	More important	Less important
Statistics Forecasting	Broad awareness of many different methods of estimation and sampling	Understanding application of qualitative and quantitative methods of forecasting

- Derivations of methods and proofs of maximum likelihood estimation
- Understanding of underlying stochastic processes
- Optimization
- Numerical methods of optimization
- Finding global optimal solutions
- Discrete event simulation
- Quick design and implementation of discrete event simulation models.

Queuing theory Applied probability Using probability theory with actual data to estimate the expected value of random variables of interest.
The theory of stochastic processes Analytical mathematical modeling.
Using numerical methods to estimate functions relating independent variables to dependent variables
Here is our proposed definition of SCM data science: SCM data science involves using SCM theory and a range of quantitative and qualitative techniques to relevant SCM challenges and outcome prediction, while taking into consideration issues with data availability and quality. The Journal of Business Logistics would be happy to publish research on this topic. Practitioners are looking for solutions, and as researchers, we ought to offer solutions, frameworks, and answers that are all based on research that is theoretically grounded. To validate, refute, and expand the concepts presented in Table 1, we require theoretically informed research. Table 1 at this point is merely speculation. We are looking for studies that focus on the skill sets that SCM data scientists require.
Some may disagree with our assertion that the SCM data scientist requires SCM theory. However, the proliferation of data and data variety highlights the significance of theory. Theory plays a crucial role in avoiding false positives3. False positives occur when relationships between variables are discovered that do not actually exist. The issue is that the likelihood of false positives increases exponentially with the number of variables, or with the use of theoretical data mining. In Figure 2, the number of variables is represented by the horizontal axis, and the number of false positives is represented by the vertical axis when the probability of a given false positive is 0.01.Theory can assist the researcher or manager in avoiding spurious decision-making by preventing them from falling prey to relationships that are only "apparent" and do not actually exist."We have found that... hypothesis-led modeling generates faster outcomes and also roots models in practical data relationships that are more broadly understood by managers," according to Barton and Court (2012) (p. 81).
The explosion of new variables that can be investigated is the result of big data, which will be discussed further down. For a very interesting discussion on this topic, see Carraway (2012).
The reason why we expect the number of false positives to grow exponentially. A key strategy for reducing false positives is to build models using the right logic and/or theory before running predictive analytics. Again, despite the fact that SCM data science is used, it must be based on theory to avoid a lot of false positives, which waste time and money. A "pure data mining approach often leads to an endless search for what the data really say," as Barton and

Court (2012) state once more (p. 81). Data science includes predictive analytics as a subset. Table 2 shows that understanding the uniqueness of predictive analytics reveals some interesting research needs. When the probability of a given false positive is 0.01, the relationship between the number of variables and the number of false positives

2.1 Dimension of Interest

1. Predictive analytics research (examples)
2. Relevant Less relevant
3. Statistics Quantitative
4. Integrating quantitative and
5. Qualitative analysis Forecasting
6. Predicting the future
7. Using forecasting techniques for evaluating what would have happened under different circumstances

Improving Lagrange Multiplier tests for autocorrelation Deriving generalized estimators of seasonal factors Optimization Minimization and maximization Assessment of the quality of the optimal solution and the ability to implement it versus near optimal solutions Use of polyhedral functions in linear programming. Discrete event simulation Quantitative analysis of a system in a stochastic setting Discrete event simulation in a business process reengineering setting Random number generation for discrete event simulation

Applied possibility Description of stochastic variable quantity, predictable values, and vagueness Data mining Exploration for designs and relationships among a large amount of variables through lots of information.

Precise analysis using artificial and unrealistic assumptions for theorems and proofs. Methods of quickly and inexpensively modeling approximate relationships between variables while still using deductive mathematical methods Proving inventory theorems that assume known, continuous demand with perfect information

An example of a research area that would be more relevant to predictive analytics and an example of a research area that would be less relevant to predictive analytics are presented in Table 2, which analyses a sample of predictive analytics-related disciplines, chooses a dimension of that subject, and contrasts prospective study areas. The difference between predictive analytics and each of these quantitative fields is indirectly shown in Table 2. In addition, it suggests potential research directions in the field of predictive analytics.

Importantly, despite its connection to numerous established quantitative methods, predictive analytics stands apart from all of them. Predictive analytics is both quantitative and qualitative, whereas statistics is quantitative. Predictive analytics adds questions about what would have occurred in the past if various conditions had been present, while forecasting is all about predicting the future. Predictive analytics also focuses on what would characterize a system that was not functioning optimally, whereas optimization is about finding the minimum or maximum of a function within constraints. Predictive analytics aims to quickly and inexpensively approximate relationships between variables while still employing deductive mathematical methods to draw conclusions, whereas analytical modeling is primarily concerned with generating mathematical axioms and then proving lemmas and theorems.

Predictive analytics and well-known quantitative fields place different emphasis on different areas, as shown by these examples.

Table 2 contains some of the topics that have been looked at, but more research in these relevant areas would help predictive analytics refine and improve supply chain decision making. In fact, logistics and SCM-related predictive analytics research is of interest to the Journal of Business Logistics. To that end, we offer the following definitions of predictive analytics for the supply chain and logistics:

What is defined here as logistics predictive analytics and SCM predictive analytics has already existed in the past, it just lacked a Direct sales, sales of distributors, Internet sales, international sales, and competitor sales

Face profiling data for shopper identification and emotion detection; eye-tracking data; customer sentiment about products purchased based on “Likes,” “Tweets,” and product reviews

Inventory in warehouses, stores, Internet stores, and a wide variety of vendors online

Not only where it is, but what is close to it, who moved it, its path to get there, and its predicted path forward; location positions that are time stamped from mobile devices name. Because the concept is becoming so widespread, using a name facilitates communication about it. If you read Table 2 with these definitions in mind, you should be able to find relevant research on SCM predictive analytics or logistics that would be especially interesting for the Journal of Business Logistics. According to Barton and Court (2012), advanced analytics are gaining in importance:

2.2 Big Data

According to McAfee and Brynjolfsson (2012), big data is unique due to the volume, variety, and velocity of the data, which is now widely available and much less expensive to access and store. There are numerous methods for volume. There are more data because the data are captured in greater detail, among other things. For instance, in addition to simply recording that a unit sold at a specific location, the time it was sold and the quantity of inventory at the time of sale are also recorded. Another example: in the past, many businesses did not keep track of daily sales by stock-keeping unit and location in order to make decisions about inventory. Furthermore, data must be collected at multiple points along long global supply chains. Furthermore, there is now a in great numbers. Some examples of big data's causes are shown in Table 3.

Big data can be used in a variety of ways in logistics and SCM, as shown in Table 4.In Table 4, each row represents a distinct type of logistics user and each column represents a crucial managerial component of business logistics. This is not meant to be an all-encompassing list of logistics users or components.

With reference to a variety of big data sources, Table 5 provides examples of research questions based on management domains in logistics and SCM.

2.3 Back to Basics

Finally, using management theory as a lens, we deliver a few examples of investigate questions that are pertinent to SCM in Table 6.

2.4 Demand for DPB Professionals

We believe that there is an increasing demand for DPB-certified professionals. We provide figures showing the rising number of Google searches for these terms as an illustration of general interest. A graph of the number of searches made by Google for various relevant terms since 2004 is shown in Figure 3.The y-axes have relative scales, with 100 representing the maximum number of searches.

Figure 3A shows that searches for "Data Scientist" and "data science" were virtually nonexistent until 2005 and 2011, respectively. We believe that the use of the terms is spreading. The proliferation of consumer sentiment data from product reviews, Tweets, and Likes on websites. It is necessary to analyze and quantify such data. There are a growing number of software companies offering algorithms for evaluating tweets and reviews.

2.5 Transportation Management

Increased demand for individuals with big data-related expertise. A graph of searches for the term "predictive analytics" is depicted in Figure 3B.Similar to searches for "Data Science," ure 3C displays a graph of searches for "Big Data" and "Supply Chain Management" as references for searches for "Predictive Analytics."As can be seen, this is the first year that "Big Data" had more searches than "Supply Chain," which essentially started in 2005 and experienced significant growth thereafter. Table 6: Examples of big data research questions that are relevant to supply chain management (SCM), stemming from management theory. How does the existence of big data affect the reduction in internal transaction costs vis-`a-vis external transaction costs and how is this affecting the size of logistics organizations and the structure of supply chains?

Resource-based view Can SCM data science be developed as a resource that is valuable, rare, in imitable, and nonsubstitutable?

2.6 Contingency Theory

How can big data and SCM data science be used by logistics managers to meet internal needs and adjust to changes in the supply chain environment

Resource dependence theory Management."We do not anticipate that SCM will surpass big data in importance;This is only shown to show how much more people are using the phrase.

Figure 3 makes it abundantly clear that interest in DPB is expanding exponentially.We believe that this trend's underlying phenomena present a number of opportunities as well as challenges for our discipline.Big data has the potential to improve supply chain design, relationship development, customer service systems, and day-to-day value-added operations management, but we have only just begun to look into the possibilities. We need to be more creative in our inquiries and make well-informed decisions about them.

III. CONCLUDING DISCUSSION

Even though the term "big data" has become popular in recent times, it has significant ramifications for our field, both in terms of opportunities and challenges for our research and teaching methods. We can easily see how predictive analytics and data science apply to SCM, but it can sometimes be more difficult to see how big data directly affects SCM.As a result, we would like to see research in the Journal of Business Logistics that clarifies the significance of DPB and big data in the supply chain.

Participation methods include:

1. Submit DPB-related manuscripts.
2. On DPB, submit Forward Thinking articles.
3. Send us a suggestion for a DPB Special Topics Forum.
4. Create a DPB Thought Leader Series.
5. Utilizing your existing research abilities and domain knowledge, begin a new DPB research project.
6. Please send us an email with your thoughts on how we can promote research on DPB outside of these categories.

REFERENCES

[1]. Barton, D., and Court, D. 2012. “Making Advanced Analytics Work for You.” Harvard Business Review 90:79–83.

[2]. Carraway, R. 2012. “Big Data, Small Bets.” Forbes. http://www.

[3]. forbes.com/sites/darden/2012/12/13/big-data-small-bets/.

[4]. Chen, H., Chiang, R., and Storey, V. 2012. “Business Intelligence and Analytics: From Big Data to Big Impact.” MIS Quarterly 36(4):1165–88.

[5]. Davenport, T., and Patil, D. 2012. “Data Scientist: The Sexiest Job of the 21st Century.” Harvard Business Review 90:70–76. Dumbill, E., Liddy, E., Stanton, J., Mueller, K., and Farnham, S. 2013. “Educating the Next Generation of Data Scientists.”

[6]. Big Data 1(1):21–27.

[7]. Mayer-Sch€onberger, V., and Cukier, K. 2013. Big Data: A Revolution That Will Transform How We Live, Work, and Think. New York: Houghton Mifflin Harcourt Publishing Company.

[8]. McAfee, A., and Brynjolfsson, E. 2012. “Big Data: The Management Revolution.” Harvard Business Review 90:60– 68.

[9]. Provost, F., and Fawcett, T. 2013. “Data Science and Its Relationship to Big Data and Data-Driven Decision Making.” Big Data 1(1):51–59.