Big Data and Supply Chain Analytics: Implications for Teaching

Jason Woldt
UW-Whitewater

Ceyhun Ozgur
Valparaiso University (USA)

Operations Management Education Review 14.
© 2020 NeilsonJournals Publishing. GALLEY PROOF

Big Data and Supply Chain Analytics: Implications for Teaching

Jason Woldt
University of Wisconsin-Oshkosh, USA

Sameer Prasad
University of Wisconsin-Whitewater, USA

Ceyhun Ozgur
Valparaiso University, Valparaiso, USA

Abstract. In environments characterized by turbulence, economic turmoil and uncertainty along with dramatic advancements in available technology and information availability, individuals with big data and supply chain analytics capability will continue to be in high demand. To reflect this new reality, universities will need to modify course content and pedagogy to meet industry needs. Based on industry data and an extensive literature review, we derive a course mapping model for big data and supply chain analytics. Specifically, we synthesize the literature on big data, structural elements of analytics, and outcomes. Using this model, we gathered 116 different supply chain analytics job descriptions posted on the world's largest job database (indeed.com). Based on these findings, we offer recommendations to align academic programs with the needs of the industry.

Keywords: big data, analytics, supply chain management, empirical.

1. Introduction

Big data has catalyzed the use of supply chain analytics to fundamentally transform the way products and services are delivered to the end consumer, especially in turbulent times. This revolution is, in part, a function of the new technologies now available to harvest data and a way to make meaningful changes to customer experiences by braiding traditional supply chain approaches with optimized solutions.

Given the transformative nature of these new technologies, it is critical that as educators also be aware and knowledgeable of these new approaches. This will better equip us to be able to shape our students' learning experiences so they can navigate and take advantage of the evolving supply chain management landscape. This includes ensuring that the supply chain analytics curriculum aligns with the new needs of the market.

A review of the literature demonstrates the value of analytics across different areas of the supply chain and verifies the connection between the use of supply chain analytics and higher levels of supply chain performance (Gunasekaran *et al.* 2017, Trkman *et al.* 2010, Wamba *et al.* 2017). Unfortunately, one of the barriers to implementing supply chain analytics strategies is the lack of trained employees (Schoenherr & Speier-Pero 2012). A recent KPMG CIO Survey (2018) found that a big data and analytics knowledge base was the most critical talent gap for organizations. LinkedIn (2018) also reported that over 151,000 data science positions were currently unfilled. To address the existing talent gap, researchers have recommended a greater presence of data science and supply chain analytics within university supply chain curriculums (Asamoah *et al.* 2017, Evans 2012, McLeod *et al.* 2017, Schoenherr & Speier-Pero 2012, Waller & Fawcett 2013a, Waller & Fawcett 2013b).

Universities have been addressing this market need by creating a slew of data science and analytics courses at the undergraduate and graduate level but are finding it difficult to deliver them given the skills gap within academia itself (Sturges 2019, Woodie 2016). Universities are also struggling to hire newly minted Ph.Ds. with a formal data science and analytics skill-set given competition from the corporate sector offering higher salaries (Woodie 2016). This shortage is encouraging universities to pivot many traditional supply chain management or operations research faculty into these data science and business analytics courses, often with little formal training or an established body of knowledge which can be easily imparted to students.

While all disciplines struggle to fill critical data scientist roles, Richey *et al.* (2016) found that filling critical roles in Supply Chain Management (SCM) is particularly difficult. Researchers have offered that big data and analytics roles in SCM require more than just a strong data science and analytical skill set to be effective. For example, Waller and Fawcett (2013b) proposed a conceptual relationship where the effectiveness of a data scientist increases exponentially with domain knowledge in logistics and SCM. They also recommended a combination of SCM theory paired with knowledge in other disciplines. Furthermore, researchers have found that the unique skill set desired for SCM data science and analytics professionals is a balanced blend of data manipulation and communication/ interpersonal skills (Schoenherr & Speier-Pero 2015, Woolsey 1981). The literature suggests that a better configuration would be to blend data science and analytics theory, concepts, and tools into an SCM-based curriculum.

In this study, we gather job postings using the search terms “supply chain analytics”. Based upon these job postings, we classify descriptors within big data, supply chain analytical methods, and operational outcomes. Using these descriptive patterns, we map out content within supply chain programs tailored to employer expectations.

For educators, such as us, it is essential that we systematically understand the salient elements of big data, supply analytics, and outcomes. Next, the relevant literature is reviewed to provide a common platform of understanding.

2. Big Data

Researchers and academics have found that big data has had a transformative impact on the economic environment in both the public and private sector across the world (Chen & Zhang 2014, Manyika *et al.* 2011). In addition, McAfee *et al.* (2012) found that supply chains characterized as data-driven were more productive and more profitable than their competitors. More specifically, Perrey *et al.* (2013) found that making business decisions using big data can increase an origination's return on investment. As such, the topic of big data and analytics has emerged as a trending topic for both researchers and practitioners in the field of SCM (Hazen *et al.* 2014, Wang *et al.* 2016). In addition, Waller and Fawcett (2013a; 2013b) found that data science, predictive analytics, and big data are essential to the development of future SCM leaders. While few disagree about the transformative impact of big data and analytics on the business and supply chain environment, the term "big data" has been used loosely in both industry and academia. To properly develop a recommended model, it is critical to first define big data within the SCM environment.

As a foundation to articulate the unique characteristics of big data, we examine the relevant literature. Researchers initially described big data using the three Vs: Volume, Variety, and Velocity (Gartner 2012, Kwon & Sim 2013, McAffee & Bryjolfsson 2012). Others have argued that defining big data using only the three Vs fails to consider important characteristics of the data such as value and veracity and have proposed a big data definition with 5 Vs including value and veracity (Wamba *et al.* 2015, White 2012). For a proposed teaching model, a more holistic definition of big data using the 5 Vs: Volume, Velocity, Variety, Veracity, and Value would be more impactful in the learning experience. In the next section, we define each of the 5 Vs as part of the big data definition.

Volume

Volume has been referred to as the amount of data or the size of the data set that is available to be collected (Chen & Zhang 2014). Over time, it has become easier and less expensive to collect data, resulting in larger quantities of available data (Wamba *et al.* 2017). For example, in 2013, IBM reported that 90 percent of all existing data was created in the last two years (Jacobson 2013). In addition, according to Domo's 6th Annual *Data Never Sleeps Report* (2018), 2.5 quintillion bytes of data are created every single day, and by 2020, it is estimated that 1.7 megabytes of data will be collected every second for every person in the world. Furthermore, this trend is likely to continue with the International Data Corporation (IDC) projecting that all of the data collected worldwide will grow from 33 zettabytes in 2018 to 175 zettabytes in 2025 (Reinsel *et al.* 2018).

Within SCM, big data has been a particularly relevant topic of discussion. In terms of volume, the increased use of Enterprise Resource Planning (ERP) systems coupled with Geographical Information Systems (GIS) provide the opportunity to capture transactional data throughout the supply chain. For example, organizations can capture and store data relevant to forecasting, purchasing, inbound logistics, inventory management, operations, order management, outbound logistics, and returns (Addo-Tenkorang & Helo 2016, Kaur & Singh 2018). Because the SCM function is either directly or indirectly responsible for the management of these activities, a big data and analytical skill set is critical. In mass manufacturing environments with broad portfolios of products and supply chains consisting of multiple nodes across multiple countries, the volume of data collected can be significant. ERP systems allow both the buying and supplying organization to capture and share this data using Electronic Data Interchange (EDI).

Velocity

Big data velocity refers to the speed at which data are generated, and the speed at which this data should be analyzed and decisions made to maximize value (Gandomi & Haider 2015). McAfee *et al.* (2012) highlighted the importance of the inverse relationship between the value of the data and time. For example, as time passes, data that is not analyzed and acted upon loses its value. Furthermore, Hofmann (2017) found that when compared to variety and volume, the presence of the velocity characteristic in a systems dynamic model has the greatest positive impact in an organization's ability to mitigate the bullwhip effect. Thus, exposing students to real-time data and the necessity of acting upon such data rapidly needs to be reinforced in the way the SCM big data and analytics curriculum is delivered.

Variety

The variety aspect of big data refers to the heterogeneity in the structure of the data (Gandomi & Haider 2015). While data is generally thought of as numbers in spreadsheets and tables, data can exist in a variety of different forms including: sensor data, text messages, social media posts, images, audio, and video (Hofmann 2017, Wamba *et al.* 2017). In fact, Cukier (2010) reported that only 5% of the data that exists is structured data in spreadsheets and databases, and most of the available data exists in an unstructured form that is not able to be easily analyzed. In SCM, Waller and Fawcett (2013b) found that because of a supply chain manager's scope of responsibility, the variety of data collected spans across functional disciplines and includes a wide variety of data. Furthermore, ERP and other technologies generate a vast amount of data from multiple sources across diverse supply chain networks. This is true for commercial, governmental, and non-profit sectors (Prasad *et al.* 2018). Often problems encountered in such contexts are quite complex, requiring SCM data scientists to assemble a variety of data into a cohesive decision support system. As students transition into the working world, they will have to make relevant decisions on which variables to consider and how to assemble the diverse forms of data into an analytical model.

Veracity

The veracity of the data is the level of reliability, correctness, and precision inherent in the data (Gandomi & Haider 2015, Richey *et al.* 2016). Researchers and practitioners alike find that the veracity of the data collected is critical to drawing correct conclusions that translate into the well-informed decisions (Aman *et al.* 2014). Similarly, LaValle *et al.* (2011) found that twenty percent of managers cited data quality as a primary barrier to implementing more robust data analytics strategies. Finally, Hazen *et al.* (2014) addressed the growing concerns over data quality issues in the supply chain and have recommended that managers take strategic steps to control the quality of the supply chain data similar to the way that organizations control the quality of products. Thus, it is critical that students have experience in questioning data veracity and take a role in creating systems to ensure data quality.

Value

The value of big data refers to its inherent usefulness or worth (LaValle *et al.* 2011). Oracle (2019) labeled big data as organizational capital but also articulated that most big data, in raw form, has very low value density. Gandomi and Haider (2015) add that relative to volume, the inherent value in raw data is low because the value is extracted after the data has been analyzed. Veeramachaneni (2016) found that organizations are often not getting the value that they expect from big data initiatives because the data is often poorly organized. Given the plethora of variables available for SCM modeling, students need to organize data, which includes making decisions about which and how many variables to use. Harris and Mehrotra (2014) also note that to build effective business models, supply chain data scientists must be integrated into the business development process at all stages of the project. Hence, shaping the value at the input phase affects the analytical transformation methods and yields the desired outcomes.

3. Supply Chain Analtyical Methods

Hazen, Skipper, Ezell, and Boone (2016) found that the value of big data can only be realized when the data is paired with analytical tools that apply math and statistical tools to summarize, interpret, and use this data to make decisions. Past research has used a three-category taxonomy to classify the analytical categories into either descriptive, predictive, and prescriptive analytics (Davenport 2013, Delen & Demirkan 2013, Souza 2014, Trkman *et al.* 2010, Wang *et al.* 2016). Each of the three categories will be defined and discussed in the subsequent sections pertaining to the supply chain application, complexity, and value-added contribution.

Descriptive Analytics

At the foundational analytics level, descriptive analytics describes or answers questions about what happened or what is happening in the supply chain (Delen & Demirkan 2013, Wang *et al.* 2016). Researchers have found that descriptive analytics enables organizations to do various levels of business reporting such as generating simple reports, scorecards, dashboards, or summarize shipment detail (Evans & Linder 2012, Souza 2014, Tiwari *et al.* 2018). In order to accomplish this task, Lade *et al.* (2017) found that organizations have increasingly incorporated software visualizations tools such as Tableau and Microsoft Business Intelligence (BI) to construct descriptive analytical tools such as histograms, multiple variable plots, and correlation analyses. Using the data and information contained in these visuals and reports, managers can then identify patterns and make associations enabling them to more positively impact business results (Tiwari *et al.* 2018; Wang *et al.* 2016). Because descriptive analytics uses traditional statistical analysis, it involves a lower level of complexity. The three-step process of reporting, interpreting, and acting upon information takes time and requires managers to correctly interpret the data before the most optimal supply chain decisions are made.

Predictive Analytics

Building upon descriptive analytics, predictive analytics is more complex and allows the organization to uncover relationships and make associations that may not be readily apparent (Ittmann 2015). At a basic level, researchers have described predictive analytics as the set of data science tools that are needed to make predictions (Barton & Court 2012, Davenport & Patil 2012). Furthermore, Waller and Fawcett (2013b) offered a formal definition of SCM predictive analytics as:

> SCM predictive analytics use both quantitative and qualitative methods to improve supply chain design and competitiveness by investigating both past

> and future levels of integration of business processes among functions or companies, as well as the associated costs and service levels (p. 80).

In a supply chain analytics review of literature, Wang *et al.* (2016), found that supply chain analytics has been used by organizations as a valuable strategy at the operational, tactical, and strategic levels. At each level, predictive analytics uses historical data to forecast supply chain activities to predict future patterns and outcomes. Specifically, researchers have used predictive analytics to identify trends likely to continue, create optimization models, mine data to uncover the relationships among variables, and create analytical and mathematical models (Ittmann 2015, Krupnik 2013, Liberatore & Luo 2010). Some notable studies utilized predictive analytics to more accurately forecast demand and optimize inventory (Downing *et al.* 2014, Farasyn *et al.* 2011), evaluate suppliers (Romano & Formentini 2012), predict supply chain disruptions (Souza 2014), identify and optimize the routing of people and goods (Drexl 2013), and select among a number of different supply chain network design configurations (Soleimani *et al.* 2014).

Compared to descriptive analytics, predictive analytics requires a more robust skill set (Sanders 2018). Methods of analysis such as time series forecasting and regression analysis can be used to establish basic relationships. However, because of big data complexity, more advanced statistical techniques have been proven to result in a higher level of precision and a greater ability to mitigate statistical biases (Wang *et al.* 2016). Predictive analytical approaches utilizing more advanced simulation procedures are more labor and cost intensive, but they also result in more value-added contribution in their increased ability to investigate processes that extend across the supply chain (Ranjan 2014).

Prescriptive Analytics

The prescriptive analytic approach builds upon descriptive and predictive analytics to not only understand what is happening now in the future but also prescribes a solution using optimization models (Ransbotham *et al.* 2015, Souza 2014). Building upon the work of Waller and Fawcett (2013b), Hahn and Packowski (2015) found that prescriptive supply chain solutions are more most often implemented at the operational and strategic levels. Today's optimization models tend to be more robust and capable of improving complex supply networks at a finer level of granularity (Bartolacci *et al.* 2012). Furthermore, Souza (2014) describes the majority of academic research and analytical activity in SCM practice as prescriptive analytics. In terms of the methodological approach utilized in SCM, Souza (2014) found that such techniques could include: analytical hierarchies, game theory, mixed integer linear programing/ nonlinear programming, and network flow algorithms.

Furthermore, Wang *et al.* (2016) described prescriptive as advanced multi-criteria decision models, optimization, and simulations.

While prescriptive analytics can be applied to any part of the business, researchers and practitioners have found significant value in converting analytical knowledge into the supply chain planning and execution processes. At the start of the supply chain, organizations have converted the predictive analytical knowledge in the forecasting process to systematically and prescriptively adjust forecasted quantities and minimize forecast error (Hassani & Silva 2015, LaValle *et al.* 2012). Furthermore, Bertsimas and Kallus (2014) found value in using machine learning and operations research to optimize inventory decisions. Ittmann (2015) also found that logistics organizations have used prescriptive analytics to select the most optimal route to reduce the delivery time and the cost of last mile deliveries. Finally, Hoffman (2017) found value in using a prescriptive analytical approach to mitigate the bullwhip effect throughout the supply chain. Given the input data and the supply chain analytical methods, organizations expect relevant outcomes to be derived from the process.

4. Outcomes

Melnyk *et al.* (2010) describe high performing supply chains as supply chains that carefully prioritize supply chain outcomes. Furthermore, Gunasekaran *et al.* (2004) state that supply chain outcomes should reflect the drivers of organizational comparative advantages.

Traditional Outcomes

While all supply chain measures cannot be generalized due to differences in strategy across organizations (Cai *et al.* 2009), there is a consensus that supply chain outcomes must be focused on an organization's ability to manage cost, quality, and lead time (Melnyk *et al.* 2010, Shepherd & Günter 2010). Also, basic textbooks in operations and supply chain management highlight three traditional conventional outcomes including cost, quality, and lead-times reflecting organizational preferences (Wisner *et al.* 2019). In today's turbulent environment, we need our students to view outcomes beyond a traditional perspective. Although operational metrics such as, costs, quality, and lead-time are still valid, students must be prepared to envision a wider range of factors that are critical to the long-term success of an organization's supply chain. In addition, today's business landscape is exponentially more competitive and organizations must ensure that they are not only poised to perform well financially, effectively manage quality, and deliver quickly at a single point in time, but organizational must critically assess outcomes that signal their ability to be successful in the future (Chavez *et al.* 2017, Gunasekaran *et al.* 2004).

While the outcomes and measures that assess current performance are often quantitative and straightforward, outcomes that are more dynamic can be both quantitative and qualitative and look beyond an organization's immediate financial performance. As such, researchers have suggested that these organizational and supply chain outcomes should reflect a balance of both financial and non-financial measures (Gunasekaran *et al.* 2001, Gunasekaran *et al.* 2004, Kaplan & Norton 1998). Specifically, within the area of supply chain management, two areas that have surfaced as progressive measures to assess organizational outcomes have been linked to a supply chain's level of innovation (Chan 2003, Shepherd & Günter 2010, Panayides & Lun 2009) and ability to mitigate risk, described as supply chain resilience (Srinivasan *et al.* 2011, Zhang *et al.* 2011). These two additional outcomes will be described as supply chain innovation and supply chain resilience and will be discussed in more detail below.

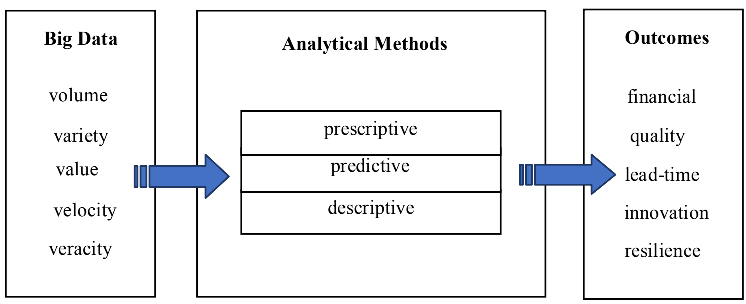
Supply Chain Innovation

Researchers have found that innovation is a key to creating an effective supply chain management strategy and can be a path to creating an organizational competitive advantage (Kwak *et al.* 2018, Wu *et al.* 2017). At an organizational level, this can be accomplished with an organization's ability to renew itself by successfully exploiting new ideas (Tidd & Bessant 2018). More specifically, Arlbjørn *et al.* (2011), define supply chain innovation as a major change in either a supply chain network, a supply chain technology, or a supply chain process. While innovation can be linked to positive outcomes in any area of the supply chain, researchers have found that supply chain innovation can be a future catalyst to improving the general supply chain outcomes of cost, quality, and service (Flint *et al.* 2005, Krabbe 2007).

Supply Chain Resilience

As the business environment and supply chains have become exponentially more complex, organizations are increasingly finding that the single largest threat to their businesses can come in the form of a supply chain disruption (Boone *et al.* 2016, Craighead *et al.* 2007, Green 2004). As discovered with the global supply chain disruptions (*i.e.* Covid-19 pandemic), shocks and uncertainties can have dramatic effects on supply chains. Big data and analytics has not only been used to identify unique business insights (Wang *et al.* 2016) but has also been used to identify and assess areas of risk and vulnerability (Souza 2014) more quickly in order to respond to such risks after they occur (Simchi-Levi *et al.* 2016). Because so many supply chain disruptions are simply unavoidable, the term supply chain resilience has emerged as an important organizational outcome. According to Hohenstein *et* *al.* (2015) this term captures a supply chain's ability to react to, cope with, and adapt to events that were not anticipated. This review found that the resilience cycle consists of four primary phases including: readiness, response, recovery, and growth (Sheffi & Rice 2005). The ability to progress through these stages quickly while utilizing minimal resources are indicators of resilient supply chains. In organizations and the field of supply chain management, identifying the elements that facilitate resilience is ripe for improvement. In most cases, complex analysis must be done to collect and analyze historical and real time data sets to ensure that organizations are optimizing critical resources and quickly executing decisions.

Figure 1: Big data, supply chain analytics methods, and outcomes – a teaching pedagogy



5. Methodology

To assess the demand for supply chain analytics professionals in the market, we executed a search within the website indeed.com. According to indeed.com, they claim to be the number one job site in the world offering a search engine to connect job seekers with job vacancies in organizations (Indeed 2020). As of 2018, the company claimed to have at least 250 million unique visitors each month (Indeed 2020). Because at least 93% of college educated job candidates have either browsed, researched, or applied to a job online (Smith 2015), we feel that an online search portal is an accurate reflection of the job market, especially for positions that are technology focused.

The indeed search engine uses basic search algorithms to filter search results based on a user's search criteria. For the purposes of this research, we used the search words "supply chain analytics" in the search query to limit our search results. In addition, our search was limited to full time positions. Part time, temporary positions, and internships were manually eliminated from the search results. Because each search is country specific, we limited our search to the United States. The search was executed on March 13th, 2020 and returned a total of 116 jobs. All 116 job postings were copied into Microsoft Excel for further analysis.

Similar to the research of Rossetti and Dooley (2010), we used downloaded job postings and applied the descriptor words to our theoretical model. Using this model, we categorized the keywords within the position description to align with our model with the categories of Big Data, Structure, and Outcomes. Each of these categories will be discussed in further detail below. The research team read, in detail, each of the 116 job postings and recorded the instances where a job required a skill set within the areas of big data (5 categories), analytical methods (3 categories) and outcomes (5 categories). Because requirements among positions varied, subgroups were created to allow for differentiation within each of the respective categories.

6. Results

An examination of some seventeen syllabi relevant to supply chain analytics programs reveals a wide variance in the range of topics covered. This variance indicates a lack of understand of industry needs. In this research, we hope distributions obtained via job posting provide a basis to create standardized curriculums tailored to industry needs.

Big Data

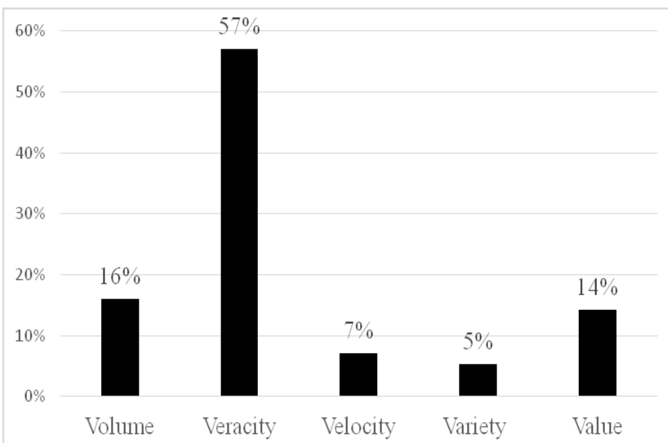
Based upon the Indeed postings, it is clear that veracity is at the forefront of industry expectations (see Figure 2), volume and value appear to be secondary, while velocity and variety are perhaps even less critical. Industry expectations relative to veracity include the ability to validate the accuracy/quality of the data, to cleanse the data, and ensure data adequacy. In most supply chain courses being taught, there is little emphasis on data veracity. As a result, there are several ways that veracity exercises could be implemented into the supply chain analytics curriculum. For example, errors in data sets could be created and students would be required to identify characteristics of accurate and inaccurate data, use statistical techniques to identify and investigate outliers, and be required to recommend process changes in data collection to ensure that future errors are minimized.

Furthermore, data veracity could be incorporated as a more significant component within supply chain capstone courses, where students are provided with real-world data, and the students need to ensure its veracity. Because data cleansing can often times be a clerical task, students could then practice creating instructions for someone else on the steps that one would need to take to cleanse the data. If working with industry partners, practitioners could share the difficulties involved in maintaining master data and offer insight into how data integrity issues could arise.

Other skill sets that might be introduced in a capstone class could revolve around value. Exercises could include sorting and categorizing data based on value to the organization. Because data is not free, students could be required to do cost-benefit analysis do determine whether or not data should be collected or should continue to be collected. Students would need to decide on which variable to including in the model building process and how to collect such data. As part of the model building process, a post hoc analysis would also be necessary to understand missing data that could have been of value.

Volume and variety might be part of a stand-alone supply analytics class, where students are given large data sets with a large number of variables and asked to manipulate the data using SAP Lumira, Tableau, or Microsoft Power BI.

Figure 2: Inclusion of big data elements in SCM analytics job vacancies



Analytical Methods

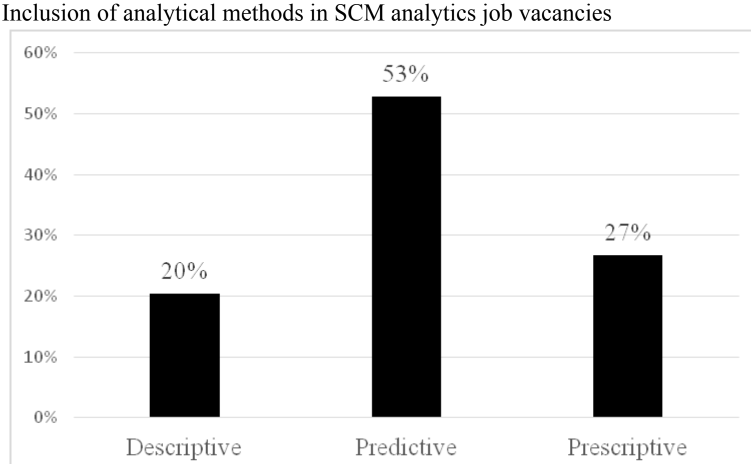
Fifty-three percent of the analytical terms fell within the predictive category. Prescriptive and descriptive methods are also quite useful (see Figure 3). It is notable that in total, 80 percent of the positions require either a predictive or a prescriptive skillset requiring a higher level of analytical knowledge and training. Perhaps organizations are finding ways to automate the descriptive work that has historically been performed by analysts.

Predictive skills include the ability to forecast, identification of anomalies, and conduct what-if and ad-hoc analyses. Most supply chain programs provide sufficient coverage on forecasting. However, the identification of anomalies is generally not part of the regular course content in supply chain management. Perhaps, within a statistics class, concepts of classification and outlier detection might be introduced to enhance students' capabilities in anomaly identification.

Prescriptive approaches noted from the job postings include the ability to create reports/recommendations to support decision making, develop/design systems, and use simulation and other optimization techniques. Simulation and other optimization techniques could be introduced within a supply chain analytics class, while the creation of reports/recommendations to support decision making and the development and design of systems could be part of a capstone class.

Descriptive methods noted in the industry job postings included visualization, creation of dashboards, and basic interpretation. Visualization potentially could be part of a statistics course, while interpretation could be part of a supply chain analytics class. Part of the capstone class could be the creation of dashboards to support Decision Support Systems (DSS).

Figure 3: Inclusion of analytical methods in SCM analytics job vacancies



7. Outcomes

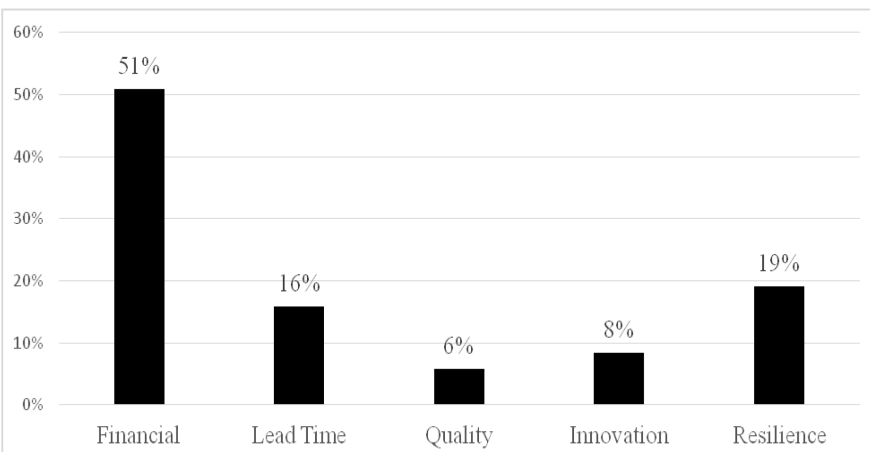
Traditional Outcomes

As we examine the deliverables specified by employers through the job postings, we noticed that the cost related terms were the largest category (see Figure 4). In the classification process, we realized the cost category needs to be more inclusive and capture metrics such as value, yield, profitability, total cost of ownership, and efficiencies. Hence, we decided to rename this category financial. Generally, such metrics are covered in the standard operations management/supply chain management courses. Perhaps, in a supply chain analytics course, such metrics could be coupled with the various descriptive, predictive, and prescriptive concepts developed in the course. Lead-time, quality, and innovation deliverables were not as significant as resilience.

Non-Traditional Outcomes

Surprisingly, supply chain resilience was the second most important category (see Figure 4), more significant than lead-time, quality, or innovation dimensions. Terms associated with supply chain resilience included operational outcomes such as agility, stability, and risk assessment/reduction. Although resilience is generally not a common topic in supply chain courses, it perhaps needs to be more integrated throughout supply chain programs. This can be integrated by having students perform stress tests on supply chains and identify areas that are most vulnerable. Resilience as a topic is expected to gain in prominence, given the shocks being propagated throughout supply chains.

Figure 4: Inclusion of outcomes in SCM analytics job vacancies



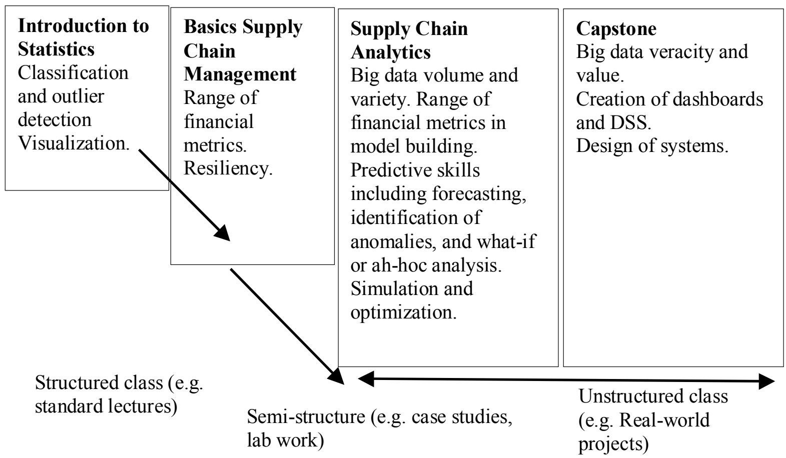
8. Discussion

The results allow us the opportunity to provide suggestions for course mapping and possible pedagogy to be employed within the respective classes in business programs at the undergraduate and graduate level.

Course Mapping

A majority of supply chain programs are anchored with an introduction to statistics and the basics of supply chain management courses. Many schools have now introduced dedicated courses in supply chain analytics, while others provide a capstone experience. In our opinion, it is possible to modify the content of these four courses and interlink learning outcomes to create program outcomes that are geared to industry expectations. Specifically, we recommend that the relevant supply chain analytics concepts are inter-dispersed among these four key classes rather than in a single course. The specific modifications in content area are shown in Figure 5.

Figure 5: Modification, linkages, and degree of structure among courses



As is common in many programs, we suggest that introduction to statistics remains a pre-requisite for the basics of supply chain management. This ensures that students have some ability to classify data and identify outliers. Furthermore, it lays the foundation to create predictive and prescriptive models in more advanced courses. We also recommend that the basics of supply chain management course serves as a pre-requisite for supply chain analytics and capstone courses. Within the basics of supply chain management, some simple predictive models can be developed, and concepts such as financial metrics and resiliency are introduced. Supply chain analytics can focus upon advanced forecasting models, big data (volume and variety), identification of anomalies, conducting what-if or ad-hoc analysis, and running simulation and optimization models. In tandem, big data (veracity and value), creation of dashboards and decision support systems (DSS), and design of systems could be part of the capstone deliverables.

Structured, Semi-Structured and Unstructured Pedagogy

Generally, introduction to statistics and basics of supply chain management can be taught via a relatively standard structured approach, especially when the class sizes tend to be quite large. However, we suggest that supply chain analytics takes a more blended or semi-structured approach, where lectures are interspersed with case-studies and lab work. On the other end of the continuum would be the unstructured nature of the real-world projects within the capstone class. As students advance through the program, they also move along the continuum closely being more immersed in the semi and unstructured processes toward the end of the program, paralleling the work the students will be expected to perform as professionals.

Students should be encouraged to find their own real-world problems, take the initiative to find and collect big data, decide upon the relevant structural tools, and deliver optimized outcomes to real-world stakeholders. In this environment, the students would be in charge and responsible for the process, while instructors would serve as guides and mentors.

The capstone experience needs to be carefully managed. In this process, it is essential that students are exposed to unstructured and uncertain situations, where they need to create the structure and mitigate supply chain risk. Many students comfortable with the traditional teaching process (standard lecture format and multiple-choice exams) would find the capstone experience somewhat difficult, with the possibility of a few student-led real-world projects possibly going awry. Because research suggests that structured teamwork tends to be more productive (Rao & Argote 2006), we suggest a certain balance along the structured-unstructured continuum.

Correspondingly, we suggest generic deliverables specified to be provided along a given timeline. Deliverables could include identifying outcomes (traditional and non-traditional metrics) as part of the problem definition process. As the course evolves, students would need to specify elements of big data that they will utilize. Toward the middle of the course timeline, students would then need to identify the analytical tools to be used. Finally, toward the end of the course, students would need to create models/systems utilizing big data.

9. Conclusion

This study is based on a growing practitioner and academic literature that identifies a pressing need for more data scientists and analytics professionals within SCM. Our initial investigation revealed that universities all over the world are answering this demand by developing courses at the undergraduate and graduate levels. We found that there is still a significant and increasing shortage of qualified faculty to teach these courses. As such, many traditional SCM faculty are pivoting to teach in this area, and a blueprint on the content and possible pedagogy to follow would be helpful.

In this study, we first conducted a review of the literature to help us understand the use of big data and analytical tools to generate positive outcomes. This literature is then used to develop a model connecting the elements of big data, supply chain analytics, and supply chain outcomes. We then compared the elements of the model to industry postings. This analysis allows us to specifically recommend course content and pedagogies interlinked within a series of relevant supply chain analytical courses.

Given an ever-changing business climate, supply chain analytics will grow in importance. Given the dramatic shifts in demand and supply patterns, companies need to have the ability to synthesize and ascertain the veracity of the big data, create the necessary predictive and prescriptive models, and orient the organizations towards a range of financial outcomes and resiliency. This research provides a roadmap in creating the next generation of supply chain managers.

References:

Addo-Tenkorang, R. & Helo, P. T. (2016), "Big data applications in operations/supply-chain management: A literature review", *Computers & Industrial Engineering*, *101*: pp. 528-543.

Aman, S., Chelmis, C., & Prasanna, V. (2014, October), "Addressing data veracity in big data applications". In *2014 IEEE International Conference on Big Data:* pp. 1-3, available at https://www.osti.gov/servlets/purl/1333162

Arlbjørn, J. S., de Haas, H., & Munksgaard, K. B. (2011), "Exploring supply chain innovation", *Logistics Research*, *3*(1): pp. 3-18.

Asamoah, D. A., Sharda, R., Hassan Zadeh, A., & Kalgotra, P. (2017), "Preparing a data scientist: A pedagogic experience in designing a big data analytics course", *Decision Sciences Journal of Innovative Education*, *15*(2): pp. 161-190.

Bartolacci, M. R., LeBlanc, L. J., Kayikci, Y., & Grossman, T. A. (2012), "Optimization modeling for logistics: options and implementations", *Journal of Business Logistics*, *33*(2): pp. 118-127.

Barton, D. & Court, D. (2012), "Making advanced analytics work for you", *Harvard Business Review*, *90*(10): pp. 78-83.

Bertsimas, D. & Kallus, N. (2014), "From predictive to prescriptive analytics", *ArXiv Preprint arXiv:1402, 5481*.

Bertsimas, D. & Kallus, N. (2020), "From predictive to prescriptive analytics", *Management Science*, *66*(3): pp. 1025-1044.

Boone, C. A., Skipper, J. B., Murfield, A., & Murfield, M. L. U. (2016), "Cargo theft in the motor carrier industry: An exploratory study", *Journal of Transportation Security*, *9*(1-2): pp 57–70.

Cai, J., Liu, X., Xiao, Z., & Liu, J. (2009), "Improving supply chain performance management: A systematic approach to analyzing iterative KPI accomplishment", *Decision Support Systems*, *46*(2): pp. 512-521.

Chan, F. T. (2003), "Performance measurement in a supply chain", *The International Journal of Advanced Manufacturing Technology*, *21*(7): pp. 534-548.

Chavez, R., Yu, W., Jacobs, M. A., & Feng, M. (2017), "Data-driven supply chains, manufacturing capability and customer satisfaction", *Production Planning & Control*, *28*(11-12): pp. 906-918.

Chen, C. P. & Zhang, C. Y. (2014), "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data", *Information Sciences*, *275*: pp. 314-347.

Craighead, C. W., Blackhurst, J., Rungtusanatham, M. J., & Handfield, R. B. (2007), "The severity of supply chain disruptions: Design characteristics and mitigation capabilities", *Decision Sciences*, *38*(1): pp. 131-156.

Cukier, K. (2010), The Economist, "Data, data everywhere: A special report on managing information", accessed June 19, 2019, available at http://www.economist.com/node/15557443.

Davenport, T. H. (2013), "Analytics 3.0.", *Harvard Business Review*, *91*(12): pp. 64-72.

Davenport, T. H. & Patil, D. J. (2012), "Data scientist", *Harvard Business Review*, *90*(5): pp. 70-76.

Delen, D. & Demirkan, H. (2013), "Data, information and analytics as services", *Decision Support Systems*, *55*(1): pp. 359-363.

DOMO (2018), "Data Never Sleeps Report 6.0", accessed June 14, 2019, available at https://www.domo.com/assets/downloads/18_domo_data-never-sleeps-6+verticals.pdf

Downing, M., Chipulu, M., Ojiako, U., & Kaparis, D. (2014), "Advanced inventory planning and forecasting solutions: A case study of the UKTLCS Chinook maintenance programme", *Production Planning & Control*, *25*(1): pp. 73-90.

Drexl, M. (2013), "Applications of the vehicle routing problem with trailers and transshipments", *European Journal of Operational Research*, *227*(2): pp. 275-283.

Evans, J. R. & Lindner, C. H. (2012), "Business analytics: the next frontier for decision sciences", *Decision Line*, *43*(2): pp. 4-6.

Farasyn, I., Humair, S., Kahn, J. I., Neale, J. J., Rosen, O., Ruark, J., & Willems, S. P. (2011), "Inventory optimization at procter & gamble: Achieving real benefits through user adoption of inventory tools", *Interfaces*, *41*(1): pp. 66-78.

Flint, D. J., Larsson, E., Gammelgaard, B., & Mentzer, J. T. (2005), "Logistics innovation: a customer valueoriented social process", *Journal of Business Logistics*, *26*(1): pp. 113-147.

Gandomi, A. & Haider, M. (2015), "Beyond the hype: Big data concepts, methods, and analytics", *International Journal of Information Management*, *35*(2): pp. 137-144.

Gartner (2012), "Big Data", accessed June 14, 2019, available at https://www.gartner.com/it-glossary/big-data/

Green, M. (2004), "Survey: Executives rank fire, disruptions top threats", *Best's Review*, *105*(5): p. 105.

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017), "Big data and predictive analytics for supply chain and organizational performance", *Journal of Business Research*, *70*: pp. 308-317.

Gunasekaran, A., Patel, C., & McGaughey, R. E. (2004), "A framework for supply chain performance measurement", *International Journal of Production Economics*, *87*(3): pp. 333-347.

Gunasekaran, A., Patel, C., & Tirtiroglu, E. (2001), "Performance measures and metrics in a supply chain environment", *International Journal of Operations & Production Management*. *21*(1/2): pp. 71-87.

Hahn, G. J. & Packowski, J. (2015), "A perspective on applications of in-memory analytics in supply chain management", *Decision Support Systems*, *76*: pp. 45-52.

Harris, J. G. & Mehrotra, V. (2014), *Getting Value from Your Data Scientists*, MIT.

Hassani, H. & Silva, E. S. (2015), "Forecasting with big data: A review", *Annals of Data Science*, *2*(1): pp. 5-19.

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014), "Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications", *International Journal of Production Economics*, *154*: pp. 72-80.

Hazen, B. T., Skipper, J. B., Ezell, J. D., & Boone, C. A. (2016), "Big Data and predictive analytics for supply chain sustainability: A theory-driven research agenda", *Computers & Industrial Engineering*, *101*: pp. 592-598.

Hofmann, E. (2017), "Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect", *International Journal of Production Research*, *55*(17): pp. 5108-5126.

Hohenstein, N., Feinsel, E., Hartman, E., & Giunipero, L. (2015), "Research on the phenomenon of supply chain resilience", *International Journal of Physical Distribution & Logistics Management*, *45*: pp. 90-117.

Indeed (2020), "About Indeed". Retrieved from: https://www.indeed.com/about

Ittmann, H. W. (2015), "The impact of big data and business analytics on supply chain management", *Journal of Transport and Supply Chain Management*, *9*(1): pp. 1-9.

Jacobson (2013), "Industry Insights: IBM Consumer Products industry Blog", accessed June 19, 2019, available at https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/

Kaplan, R. S. & Norton, D. P. (1998), "Putting the balanced scorecard to work", *The Economic Impact of Knowledge*, *27*(4): pp. 315-324.

Kaur, H. & Singh, S. P. (2018), "Heuristic modeling for sustainable procurement and logistics in a supply chain using big data", *Computers & Operations Research*, *98*: pp. 301-321.

KPMG (2018), "Harvey Nash/KPMG CIO survey 2018", accessed June 14, 2019, available at https://assets.kpmg/content/dam/kpmg/be/pdf/2018/07/CIO_Survey_2018_Harvey_Nash_report.pdf

Krupnik, Y. (2013), "7 ways predictive analytics helps retailers manage suppliers", *Predictive Analytics Times,* 26 August 2013, accessed June 14, 2019, available at http:// www.predictiveanalyticsworld.com/patimes/7-ways-predictive-analyticshelps-retailers-manage-suppliers/

Kwak, D. W., Seo, Y. J., & Mason, R. (2018), "Investigating the relationship between supply chain innovation, risk management capabilities and competitive advantage in global supply chains", *International Journal of Operations & Production Management*, *38*(1): pp. 2-21.

Kwon, O. & Sim, J. M. (2013), "Effects of data set features on the performances of classification algorithms", *Expert Systems with Applications*, *40*(5): pp. 1847-1857.

Lade, P., Ghosh, R., & Srinivasan, S. (2017), "Manufacturing analytics and industrial internet of things", *IEEE Intelligent Systems*, *32*(3): pp. 74-79.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011), "Big data, analytics and the path from insights to value", *MIT Sloan Management Review*, *52*(2): pp. 21.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011), "Big data, analytics and the path from insights to value", *MIT Sloan Management Review*, *52*(2): pp. 21.

Liberatore, M. J. & Luo, W. (2010), "The analytics movement: Implications for operations research", *Interfaces*, *40*(4): pp. 313-324.

LinkedIn (2018), "Workforce Report United States August 2018", accessed June 19, 2019, available at https://economicgraph.linkedin.com/resources/linkedin-workforce-report-august-2018?trk=lilblog_08-20-18_data-scientists-America-great_tl&cid=70132000001AyziAAC

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011), "Big Data: The Next Frontier for Innovation, Competition, and Productivity", accessed June 19, 2019, from McKinsey Global Institute available at https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012), "Big data: the management revolution", *Harvard Business Review*, *90*(10): pp. 60-68.

McLeod, A. J., Bliemel, M., & Jones, N. (2017), "Examining the adoption of big data and analytics curriculum", *Business Process Management Journal*, *23*(3): pp. 506-517.

Melnyk, S. A., Davis, E. W., Spekman, R. E., & Sandor, J. (2010), "Outcome-driven supply chains", *MIT Sloan Management Review*, *51*(2): pp. 33-38.

Oracle (2019), "What is Big Data?", accessed June 19, 2019, available at https://www.oracle.com/ca-en/big-data/guide/what-is-big-data.html

Panayides, P. M. & Lun, Y. V. (2009), "The impact of trust on innovativeness and supply chain performance", *International Journal of Production Economics*, *122*(1): pp. 35-46.

Perrey, J., Spillecke, D., & Umblijs, A. (2013), "Smart analytics: How marketing drives short-term and long-term growth", *McKinsey Quarterly*, 00425-3.

Prasad, S., Zakaria, R., & Altay, N. (2018), "Big data in humanitarian supply chain networks: A resource dependence perspective", *Annals of Operations Research*, *270*(1-2): pp. 383-413.

Ranjan, R. (2014), "Modeling and simulation in performance optimization of big data processing frameworks", *IEEE Cloud Computing*, *1*(4): pp. 14-19.

Ransbotham, S., Kiron, D., & Prentice, P. K. (2015), "Minding the analytics gap", *MIT Sloan Management Review*, *56*: pp. 63-68.

Rao, R. D. & Argote, L. (2006), "Organizational learning and forgetting: The effects of turnover and structure", *European Management Review*, *3*: pp. 77-85.

Reinsel, D., Gantz, J., & Rydning, J. (2018), "The digitization of the world: from edge to core", *IDC White Paper Doc# US44413318*. accessed June 19, 2019, available at https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

Richey Jr., R. G., Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016), "A global exploration of big data in the supply chain", *International Journal of Physical Distribution & Logistics Management*, *46*(8): pp. 710-739.

Romano, P. & Formentini, M. (2012), "Designing and implementing open book accounting in buyer–supplier dyads: A framework for supplier selection and motivation", *International Journal of Production Economics*, *137*(1): pp. 68-83.

Rossetti, C. L. & Dooley, K. J. (2010), "Job types in the supply chain management profession", *Journal of Supply Chain Management*, *46*(3): pp. 40-56.
Sanders, N. (2018), *Supply Chain Analytics Understanding the Business Context, 1st edition*, Prospect Press Inc., Burlington, VT.
Schoenherr, T. & Speier-Pero, C. (2015), "Data science, predictive analytics, and big data in supply chain management: Current state and future potential", *Journal of Business Logistics*, *36*(1): pp. 120-132.
Sheffi, Y. & Rice Jr, J. B. (2005), "A supply chain view of the resilient enterprise", *MIT Sloan Management Review*, *47*(1): pp 41-48
Shepherd, C. & Günter, H. (2010), "Measuring supply chain performance: current research and future directions". In *Behavioral Operations in Planning and Scheduling* (pp. 105-121). Springer, Berlin, Heidelberg.
Simchi-Levi, D., Schmidt, W., Wei, Y., Zhang, P. Y., Combs, K., Ge, Y., & Zhang, D. (2015), "Identifying risks and mitigating disruptions in the automotive supply chain", *Interfaces*, *45*(5): pp. 375-390.
Smith, Aaron (2015), "Searching for Work in the Digital Era Pew Research Report". Retrieved from: https://www.pewresearch.org/internet/2015/11/19/searching-for-work-in-the-digital-era/
Soleimani, H., Seyyed-Esfahani, M., & Kannan, G. (2014), "Incorporating risk measures in closed-loop supply chain network design", *International Journal of Production Research*, *52*(6): pp. 1843-1867.
Souza, G. C. (2014), "Supply chain analytics", *Business Horizons*, *57*(5): pp. 595-605.
Srinivasan, M., Mukherjee, D., & Gaur, A. S. (2011), "Buyer–supplier partnership quality and supply chain performance: Moderating role of risks, and environmental uncertainty", *European Management Journal*, *29*(4): pp. 260-271.
Sturges, L. (2019), "Plugging the data science gap in academia and industry", accessed June 19, 2019, available at https://www.elsevier.com/connect/plugging-the-data-science-talent-gap-in-academia-and-industry
Tidd, J. & Bessant, J. R. (2018), *Managing Innovation: Integrating Technological, Market and Organizational Change*, John Wiley & Sons.
Tiwari, S., Wee, H. M., & Daryanto, Y. (2018), "Big data analytics in supply chain management between 2010 and 2016: Insights to industries", *Computers & Industrial Engineering*, *115*: pp. 319-330.
Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010), "The impact of business analytics on supply chain performance", *Decision Support Systems*, *49*(3): pp. 318-327.
Veeramachaneni, K. (2016), "Why you're not getting value from your data science", *Harvard Business Review*, *7*.
Waller, M. A. & Fawcett, S. E. (2013a), "Click here for a data scientist: Big data, predictive analytics, and theory development in the era of a maker movement supply chain", *Journal of Business Logistics*, *34*(4): pp. 249-252.
Waller, M. A. & Fawcett, S. E. (2013b), "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management", *Journal of Business Logistics*, *34*(2): pp. 77-84.
Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015), "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study", *International Journal of Production Economics*, *165*: pp. 234-246.
Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017), "Big data analytics and firm performance: Effects of dynamic capabilities", *Journal of Business Research*, *70*: pp. 356-365.
Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016), "Big data analytics in logistics and supply chain management: Certain investigations for research and applications", *International Journal of Production Economics*, *176*: pp. 98-110.
White, M. (2012), "Digital workplaces: Vision and reality", *Business Information Review*, *29*(4): pp. 205-214.

Wisner, J. D., Tan, K. C., & Leong, G. K. (2019), *Principles of Supply Chain Management: A Balanced Approach*, Cengage Learning.

Woodie, A. (2016), “Datanomi: Finding long term solutions to the data scientist shortage”, accessed June 19, 2019, available at https://www.datanami.com/2016/03/28/finding-long-term-solutions-data-scientist-shortage/

Woolsey, G. (1981), “Yet another essay on effective communication, or on learning where to begin, how to listen, and when to stop”, *Interfaces 11*(1): pp.10-12.

Wu, K. J., Tseng, M. L., Chiu, A. S., & Lim, M. K. (2017), “Achieving competitive advantage through supply chain agility under uncertainty: A novel multi-criteria decision-making structure”, *International Journal of Production Economics*, *190*: pp. 96-107.

Zhang, L., Wang, S., Li, F., Wang, H., Wang, L., & Tan, W. (2011), “A few measures for ensuring supply chain quality”, *International Journal of Production Research*, *49*(1): pp. 87-97.