RESEARCH PAPER

Leveraging Data Science to Enhance your Supply Chain and Improve your Company's Performance

Amrou Awaysheh
Assistant Professor of Operations Management
Director, Global Supply Chain Enterprise, (gSCIE)
Kelley School of Business Indianapolis

ASCM ASSOCIATION FOR SUPPLY CHAIN MANAGEMENT

INDIANA UNIVERSITY

Big Data, Artificial Intelligence, Machine Learning, Data Science, among others are new words that we have been hearing regularly recently. These are all part of the new data revolution that is changing business everywhere. This digitization of the supply chain will have a profound impact on supply chains and how they perform. Understanding data science will enable supply chain managers to better understand its impact on their firm and how they can leverage to enhance performance and profitability.

We are currently going through the fourth industrial revolution or Industry 4.0. This revolution driven by the availability of various enterprise data sources, enable managers to make decisions based on real-time data and analytics. This revolution will transform supply chains and enable managers to make informed decisions leveraging advanced analytics that are based on real-time data.

The goal of this research is to introduce various elements of supply chain digitization. We will introduce the various elements of digitization on the supply chain and go through some specific examples. We will then discuss a strategy for how they can be rolled out effectively in your organization.

DATA

Data can exist in many different formats. Many of us think of data as spreadsheets in Microsoft Excel. We understand its limits and understand what kind of analysis we can run or models we can build. However, the tools that are available now are far more advanced than the tools we have been using since the 1990's.

Industry 4.0 enables us to gather large amounts of data not previously possible in operations and supply chains. Supply chain managers can make better decisions when they have access to all the data and have a global understanding of the issue that they are facing. This data can be generated from several sources: individuals, machines, or sensors.

This has provided the ability of data being available that simply was not fathomable a few years ago. This is called Big data (BD). In order to help describe big data, we can think of it as the five Vs—Volume, Velocity, Variety, Veracity, and Value.

Volume: There is a very large number of cases and variables per case. The size of BD datasets often exceeds the capability of conventional analysis software, like SPSS and SAS. For instance, when we look at an Excel spreadsheet, the number of rows can be considered the Volume.

Velocity: BD is generated, processed and made available for use at a high rate of speed. For example, Walmart processes more than 2.5 petabytes of transactional data per hour. So this is how frequently a new row is generated. Imagine analyzing a spreadsheet that generates 10,000 new rows every second.

Variety: BD is stored in several ways. Data can measure multiple things as well as provide additional characteristics for each data point. Thinking in Excel terms, Variety is the number of columns that each row contains.

Veracity: The veracity or reliability of BD may be questionable, with construct validity implications. Even with an Industrial Internet of Things (IIoT) device: How is it calibrated? When was it last calibrated? How accurate is the calibration?
Value: What is the value of these different kinds of data, at some point there is diminishing returns where the cost of gathering and storing the data outweighs the value that can be extracted from that data.

In addition to Big data, there are other types of data that managers need to be aware of. Thick Data exists in formats that are difficult to analyze using conventional methods. For instance, meeting notes from managers, manuals for equipment, or daily report sheets from machine operators. All these could be hundreds if not thousands of pages in length. Yet this information is not easy to go through or sort even though there is a lot of value in it. Thick data is typically data that is unstructured and not organized in relational databases. We typically associate data to be in relational datasets to be organized in rows and columns.

A Data Lake is a system that a company has in place to store all its data. This system can store all types of data that the company has regardless of the type. These databases are also typically linked together so the user can understand how to leverage more than one dataset. Data lakes can also showcase all the data that the company has to be analyzed. Employees can then access this data through a single source. Data lakes are relatively new, and many companies are spending lots of time and energy in setting up data lakes. However, once this data lake becomes available the value for the company can be substantial.

Finally, Dark data is data that exist within an organization, yet it is not examined or analyzed and thus managers do not use it to gain any insights into the firm's operations.

Sometimes managers themselves don't know such data exists in their systems.
Such data typically does not exist in perfectly formatted BD datasets ready for analysis and interpretation. This is best thought of as an iceberg or an unknown-unknown. You do not really know what is underneath or how much ice there is if you are only looking at what is above the water.

DATA ANALYSIS

The next step is understanding the ways this data can be analyzed. There are three major categories of data that we need to consider. Descriptive Analytics provides a description of what has happened, Predictive Analytics helps describe what could happen, and Prescriptive Analytics helps managers decide what should happen.

Descriptive Analytics are a type of analytics the are traditionally conducted to describe what has happened. When most of us think of analytics now, we are probably thinking about descriptive analytics. These types of analysis, identify correlations, trends and patterns in the data. By providing insight into what previously has happened, descriptive analytics allows mangers to learn from the past. By understanding what has happened in the past. Managers can then base their decisions on what has succeeded in the past.

Predictive Analytics help managers determine what could happen under various scenarios. Forecasting, simulations, and likelihood analysis are all types of prescriptive analytics. Predictive analytics can provide managers with possible outcomes and what-if scenarios. Predictive analytics brings together this insight by linking different types of datasets. To extract the most value. Mangers can then make decisions with a strong understanding of what they expect the outcome to be.

Prescriptive Analytics are the most advanced types of analytics. These types of analytic use optimizations, machine learning, and artificial intelligence to detail what should be done. On one hand, with predictive analytics we gain an understanding of the impact of each of the different types of choices. On the other hand, prescriptive analytics helps come up with the different options that are available.

	Descriptive Analytics	Predictive Analytics	Prescriptive Analytics
	What has happened?	What could happen?	What should happen?
What kind of Supply Chain questions can we ask?	How has Supplier A been performing over the past year?	How will Supplier B perform if they are asked to manufacture 500 more units?	Which supplier should we select to manufacture our new product?
What kind of data do we need?	Supplier A's performance	The performance of all the firm's suppliers	The performance of all of the firm's suppliers as well as additional background information on other potential suppliers and secondary data sets.
Types of analysis	Regression & correlation	Forecasting and simulations	Machine learning and Artificial Intelligence

LEVERAGING DATA SCIENCE

Implementing Data-based decisions requires a fundamental change in how companies think about their data. Companies need to understand how they can leverage this data for their success. Data that currently exists in silos needs to be brought tougher. Companies need to connect data sources, analytics and data scientists to understand the big picture.

It is important to understand the overall business strategy and goals. We can't just look at data science and results in a vacuum. We need to be able to relate the results and the implications to the business and the changes needed to incorporate new recommendations. Business decisions that are based on sound analysis that examines good data are much better decisions.

A Gartner study identified that one important element for an organization to get insights from its data is to create citizen data scientists[1]. A Citizen Data Scientist is an employee who has existing functional expertise and is then upskilled to gain knowledge in data science. This requires data democratization and empowering employees to expand their skill set. Providing them with access to data as well as tools and techniques to analyze this data. This will then allow that individual to leverage the data and derive insights from data to enable making data-sound decisions that align with corporate strategy. These Citizen Data Scientists will not reach to the level of expertise of data scientists. But the Citizen Data Scientists can be equipped with tools and techniques to allow them to derive substantial insights.

Once these individuals understand the basics of data science, they will then be able to derive insights from this data. This can be accomplished by asking the right questions and leveraging the existing data to develop insights. These insights can then help solve business problems and will enable the firm to make faster decisions that are grounded in fact. Firms can also have teams of data scientists. Members of this team would be functional experts in the business and can be deployed to examine different issues facing the firm. These teams could be tasked with solving different and more complex problems as they are able to run more advanced analytics.

PRACTICAL EXAMPLES

While leveraging employees to extract the most value out of a company's data is ideal. Sometimes external consultants can shed new light by asking different questions or analyzing different types of data. I have been fortunate enough to have worked with several companies to help explore the opportunities to enhance performance through data science. Here are some examples of previous projects I've been involved in with some Fortune 500 firms. The goal of this to help provide some context of what is possible with data science.

Reducing Waste and Reducing Costs

As firms try to improve their sustainability performance, one element that they can focus on is reducing what they send to landfill. As an example, a large industrial goods manufacturer wanted to reduce the waste that they send to landfill. We worked together on a project to decrease the waste sent to landfill from its 100 manufacturing facilities and over 250 offices, warehouses, and distribution centers. This resulted in over 96 million pounds of waste diverted from landfill on an annual basis. By recycling and recovering energy from this waste, the manufacturer was able to save over $36 million annually. Data science enables us to identify the sites that were leaders as well as those that were laggards. Our analysis allowed us to control for location in addition to the type of facility, helping standardize the current performance of sites to rank their performance. We then were able to take best practices and share them across all the sites. We were also able to devote the time needed on the sites that needed substantial improvement.

Leveraging Big Data to Reduce Energy Consumption

A consumer goods manufacturer that I've worked with had installed Industrial Internet of Things (IIoT) sensors throughout their manufacturing facilities. These sensors provide data every minute from various locations throughout the factory. With many factories spread throughout the world these sensors were producing over 15,000 data points every minute, and over 21 million new data points every day. Managers were drowning in data and not sure how to leverage this data or how to proceed. By leveraging this Big Data we were able to identify certain facilities that were consuming the same amount of energy during non-working periods (evenings and weekends) as they do when there was full production. These facilities were effectively wasting energy, just like energy is wasted by leaving all the lights on in the house when no one is home. We developed a strategy to help each site understand when they were effectively wasting energy so they could shutdown equipment and reduce wasted energy. Over 4 years this resulted in over $186 million in energy savings overall.

[1] https://www.gartner.com/en/documents/3534848

Predicting Raw Material Arrival

Many companies are trying to reduce the amount of raw material inventory they have by introducing lean operations principles. A large durable goods manufacturer firm was having trouble with arrivals of raw materials to their manufacturing facilities. All the company's factories operated using Just in Time (JIT), meaning that material arrived when it was needed, and very little was stored in inventory. However, there were problems with some inbound shipments arriving late. This was causing manufacturing lines to shutdown causing substantial losses of over $5,000 per minute. The company installed trackers on incoming trucks that would provide the exact location of the truck. These trackers provide the location and additional information about the truck. This information was updated every 15 seconds for over the 2,000 incoming trucks. The amount of data that this generated was enormous. While each incoming shipment could be identified, the firm could not understand when the truck was supposed to arrive. We developed a model that allows the firm transparency into the incoming shipments. This model was able to help the firm identify incoming shipments that were going to be delayed by leveraging the tracker information, carrier performance, as well as traffic patterns. This enabled us to predict when a truck would arrive and understand the impact of that delay on the production schedule. The algorithm would then produce a revised production schedule that would rearrange production schedules to eliminate line shutdowns while still ensuring that all required orders are filled. The firm was able to realize annual savings over $40 million annually by eliminating line shutdowns.

Using Data Science to enhance your Competitive Advantage

Leveraging Data Science can help your firm compete and help enhance your supply chain's performance. Understanding data and using it to your advantage empowers supply chain managers to improve their firm's performance and profitability. In summary, here are some things to keep in mind as you integrate Data Science in your supply chain. First, it is important that your company understand the data that it generates and that it has available for analysis. Secondly, the data needs to be formatted correctly so it can be analyzed effectively. Thirdly, the data needs to be accessible so all pertinent employees in the organization can access and analyze it. Fourthly, employees need to be given tools and training to allow them to analyze the data. Finally, managers need to leverage the insight and value gleaned from the data in their managerial decision making.

ABOUT ASCM

The Association for Supply Chain Management (ASCM) is the global leader in supply chain organizational transformation, innovation and leadership. As the largest nonprofit association for supply chain, ASCM is an unbiased partner, connecting companies around the world to the newest thought leadership on all aspects of supply chain. ASCM is built on a foundation of APICS certification and training spanning 60 years. Now, ASCM is driving innovation in the industry with new products, services and partnerships that enable companies to further optimize their supply chains, secure their competitive advantage and positively influence their bottom lines.
For more information, visit ascm.org.

ASSOCIATION FOR SUPPLY CHAIN MANAGEMENT