

UNIT 2

Defining the supply chain-The supply chain as a complex System-Supply chain models and modelling-The use of computer modeling in the supply chain. Data in supply chain-data analytics Methodology-Data type-small and big data in supply Chain-Challenges in working with data.

What is a supply chain?

A supply chain is the network of all the individuals, organizations, resources, activities and technology involved in the creation and sale of a product. A supply chain encompasses everything from the delivery of source materials from the supplier to the manufacturer through to its eventual delivery to the end user. The supply chain segment involved with getting the finished product from the manufacturer to the consumer is known as the [distribution channel](#).

Steps in the supply chain

The fundamental steps of a supply chain in order are as follows:

- Sourcing raw materials.
- Refining those materials into basic parts.
- Combining those basic parts to create a product.
- Order fulfillment/Sales.
- Product delivery.
- Customer support and return services.



Supply

chain models

There are several common supply chain business models that supply chains fit into. The models have two main focuses: responsiveness and efficiency. Each model strives for some combination of both but approaches those goals differently. In addition, models tend to favor one over the other. Organizations can evaluate the value proposition of each in relation to their goals and constraints, and choose which suits them best.

The model types are:

- Continuous flow model -- works best for mature industries with a degree of stability.
- Agile model -- works best for industries with unpredictable demand and products that are made to order.
- Fast chain model -- works best for products with a short lifecycle, such as fashion items.
- Flexible model -- works best for industries with a level of stability and a few relatively predictable demand peaks.
- Custom configured model -- focuses on customizing.
- Efficient chain model -- works best for highly competitive markets in which pricing plays a large part.

SUPPLY chain as a complex system

The **supply chain as a complex system** is a conceptual framework that emphasizes the interconnected, dynamic, and adaptive nature of supply chains. Instead of seeing the supply chain as a linear process—from supplier to manufacturer to distributor to retailer to customer—it is viewed as a **network of interacting entities**, where changes in one part can ripple unpredictably across the whole system.

Here's a breakdown of what makes the supply chain a **complex system**:

1. Interconnectedness

- Every node (supplier, factory, logistics provider, etc.) is connected to others.
- A disruption (e.g., a factory shutdown, port congestion, political event) can have widespread effects.

- These connections often form **non-linear** relationships—small changes can have large impacts.

2. Feedback Loops

- Decisions in one area create feedback that affects other areas.
 - Example: A retailer increases inventory due to demand spikes → upstream suppliers react by increasing production → possible overstock when demand normalizes.
- **Bullwhip effect**: minor fluctuations in demand at the consumer level get amplified upstream.

3. Adaptation and Learning

- Supply chains evolve as participants adapt to:
 - Changing market conditions
 - Technological innovations
 - Environmental shifts (climate, regulations, geopolitics)
- Companies use data analytics and AI to dynamically reconfigure logistics, sourcing, and production.

4. Emergence

- System-level behaviors emerge that aren't predictable from individual parts.
 - Example: A resilient supply chain may emerge from decentralized decision-making, even without a central controller.
 - Conversely, systemic fragility may arise unexpectedly.

5. Non-linear Dynamics & Chaos

- Outcomes are not always proportional to inputs.
- Complex supply chains can exhibit **chaotic behavior**, where tiny variations in initial conditions (e.g., timing, demand) lead to dramatically different outcomes.

6. Multi-scale Operation

- Operates across different **geographical scales** (local to global) and **time scales** (seconds for decisions in automated systems to years for capacity building).
- Each layer of scale has its own dynamics that interact with others.

7. Decentralized Control

- No single entity controls the entire supply chain.

- Coordination depends on contracts, trust, shared platforms, and information exchange (like ERP or blockchain systems).

8. Vulnerability and Resilience

- Complex systems are inherently vulnerable to **shocks** but can also be designed to **absorb and adapt**.
 - Resilience strategies: diversification, nearshoring, inventory buffers, digital twins.

Supply chain modeling:

- This is an analytical process that seeks to understand a supply chain's structure and optimize its performance. It involves using a digital representation of a supply chain to simulate hypothetical scenarios to understand how changes in production, transportation, inventory management, and distribution are likely to impact efficiency, costs and margins, customer satisfaction, and a company's ability to meet its overall business goals.

6 types of supply chain models

Supply chain models are frameworks businesses use to manage the flow of goods, information, and finances from suppliers to customers. There are six commonly cited models, but many companies may blend multiple approaches depending on their business goals and priorities.

Supply chain management is usually broken down into six separate categories. All six supply chain models have the same goal, rely on similar components, and fit into one of two categories: efficiency or responsiveness. Each supply chain model can have a primary focus on either. Understanding these different supply chain models makes it easier to manage and protect the supply chain within your organization

The Continuous Flow Model

The continuous flow model or continuous model is best for mature industries that produce a uniform set of goods and expect demand stability from their market. Goods are in a continuous flow, and the entire chain is built for a continued, scheduled delivery of goods.

A good example of the continuous flow model would be PepsiCo. This mature company has a large customer base with little variety in demand, regardless of market conditions. Logistics has been designed so that a continuous flow of ingredients is received to produce its products and restock its vendors continuously.

The Fast Chain Model

The [fast chain model](#) is used by businesses that manufacture finished products with a short market lifecycle. It's a suitable model for businesses where products are changed frequently and delivered to the market quickly before customer demand drops, and the trend loses relevance, like budget fashion industries.

Adidas is a prime example of this supply chain model. They set up new delivery systems and supply lines regularly and create new shoes and products as quickly as possible to sell apparel to meet high demand peaks before the trend changes. End-to-end supply chain efficiency is extremely important to managers in this field.

The Efficient Chain Model

The [efficient chain model](#) is used by supply chain managers operating in highly competitive markets. They must strive for high overall equipment efficiency at all times to retain their competitive advantage, which is why they opt for the most cost-effective model when determining their supply chain strategies. Emphasis is placed on proper inventory management and maximizing the outputs of the labor force and production processes.

Nestlé is a good example of a brand that requires an efficient model. Because their margins are thin and the brand operates in very competitive markets, they must reduce costs and keep vendors well-stocked at all times.

The Agile Model

The [agile model](#) demands responsive supply chains to sell trendy products during peak demand. This model tracks market demand in real-time, co-manages its inventory, utilizes collaborative product designs, and emphasizes end-to-end efficiency. The agile model is suited to businesses that sell specialty products that require extra care in transportation or design, like luxury fashion items.

Unlike the efficient chain model that requires high volumes, the agile model is only profitable until a threshold of volume is met. The market is highly sensitive and subject to unpredictable demand or demand variation in this case. The supply chain network has to remain flexible and alert at all times. [ZARA](#) is an example of a brand that follows the agile model. Designers are quick to spot trends, draw up new designs and ensure that the entire chain responds as quickly as possible.

The Custom-Configured Model

The [custom-configured model](#) combines the agile and continuous flow models. It's used in industries where multiple custom configurations are required, e.g. brands like VANS that invite customers to log on to their site and design their own

sneakers. The market remains fairly consistent, with a few predictable demand peaks throughout the year, but the company remains flexible just in case a particular customization falls out of vogue.

Providing custom configurations means that the assembly and production plants must change their setup regularly, hence the need for agility. This model is sometimes used for prototype design and small batches, e.g. limited release products. Customer satisfaction remains one of the most important business objectives for this model.

The Flexible Model

The [flexible model](#) allows businesses to meet peaks of high demand and long periods of low demand. Suppliers are diversified and hired on a seasonal contract to ensure a steady flow of raw materials as required, and factories use automation wherever possible.

Staples is a good example of a company that follows a flexible supply chain model. The company's supply chain strategies reflect the high seasonal demand for school supplies after the summer vacation and ensure that stores are well-stocked during the rest of the year when demand is lower.

What Is Computer Modeling in the Supply Chain?

Computer modeling involves using software to simulate, analyze, and optimize the flow of goods, information, and finances in a supply chain. It creates a digital representation of real-world supply chain operations, enabling businesses to test scenarios without disrupting actual operations.

Applications in Supply Chain Management

1. Demand Forecasting

- o Uses historical data and statistical models (e.g., time series, regression, machine learning) to predict future product demand.
- o Helps optimize inventory levels and reduce stockouts or overstocking.

2. Inventory Management

- o Models simulate inventory levels under various conditions to optimize reorder points, safety stock, and lead times.
- o Example: EOQ (Economic Order Quantity) and just-in-time (JIT) strategies.

3. Network Design

- o Optimizes the layout and number of warehouses, factories, and distribution centers.
- o Considers location, transportation cost, service levels, and production capacities.

4. **Transportation and Logistics**

- o Models routes and delivery schedules to reduce fuel costs and delivery times.
- o Applications include vehicle routing problem (VRP) and transportation management systems (TMS).

5. **Production Planning and Scheduling**

- o Simulates production runs, resource allocation, and maintenance schedules to optimize throughput.
- o Reduces downtime and increases productivity.

6. **Risk Management**

- o Models help predict disruptions (e.g., supplier failures, natural disasters) and assess their impact.
- o Supports development of contingency plans.

Benefits

- **Informed Decision-Making:** Simulations help evaluate “what-if” scenarios without real-world risk.
- **Cost Reduction:** Identifies inefficiencies and reduces waste in inventory, logistics, and production.
- **Improved Responsiveness:** Enhances the ability to respond to demand changes or supply chain disruptions.
- **Optimization:** Finds the best combination of supply chain components for cost, speed, and reliability.

Popular Tools and Technologies

- **AnyLogic, Arena, Simul8** – for discrete-event simulation.
- **SAP SCM, Oracle SCM Cloud, Kinaxis RapidResponse** – integrated planning tools.
- **Python/R** with libraries like **SimPy, PuLP, or SciPy** – for custom modeling and optimization.

- **AI/ML Platforms** – for predictive analytics and adaptive supply chain models.

Real-World Example

Amazon uses complex simulation models to:

- Optimize warehouse operations and picker paths.
- Forecast demand down to zip codes.
- Automate reordering and dynamically reroute deliveries.

Data in supply chain

Data in supply chain" refers to all the information collected, processed, and used throughout the entire supply chain network — from raw material sourcing to end customer delivery. This data plays a crucial role in improving efficiency, reducing costs, and enhancing decision-making.

Key Types of Data in Supply Chain:

1. Inventory Data

- o Stock levels
- o Location of goods
- o Inventory turnover rates

2. Order Data

- o Purchase orders
- o Sales orders
- o Order status and tracking

3. Supplier Data

- o Lead times
- o Performance metrics (on-time delivery, quality)
- o Costs and contract terms

4. Transportation & Logistics Data

- o Shipping schedules
- o Carrier performance
- o Freight costs
- o Real-time tracking (GPS)

5. Customer Data

- o Demand forecasts
- o Purchase history
- o Preferences and feedback

6. Production Data

- o Machine performance
- o Output rates
- o Downtime and maintenance records

7. Financial Data

- o Cost of goods sold (COGS)
- o Profit margins
- o Payment terms and transactions

8. Sustainability Data

- o Carbon emissions
- o Waste levels
- o Energy usage

Data Analytics Methodology

Data Analytics Methodology refers to the structured process used to collect, analyze, interpret, and visualize data in order to support decision-making and generate insights. Below is a commonly accepted methodology used in data analytics projects, often aligned with frameworks like CRISP-DM (Cross-Industry Standard Process for Data Mining) or modern agile approaches.

Step 1: Data identification

This can be considered as the initial step in the process as time and resources should be allocated to identify what potential data may be available for the task at hand. At this step, knowledge of what constitutes the data to be collected is important, as well as what data is available. The sources from where the data needs to be collected need to be identified at this step. Data may or may not be available, and particular managerial tasks need to be put in place for identifying the data sources. There are situations, where there may be the possibility for a particular aspect of data to be collected; however, the organization may not have collected this type of data, or they may have collected it but not necessarily put in the performance measurement system. There are other situations where large sets of data are being generated but they do not end up into any measurement system.

Identifying the required data may be a lengthy process, but cannot be ignored as it may affect the next steps in the data analysis process.

Step 2: Data collection

This is a very important process that needs to take place to ensure that the outcomes of the analysis have the potential to generate value to the organization and to the supply chain. Particular methods have been used by researchers and practitioners to collect data (via questionnaires, surveys, generated by different systems or simulations and saved within a data warehouse system and others). Each individual method used to collect data will require a process of verification and validation applied to the data collection instrument used. An analyst needs to ensure that the data they want to collect is the actual data intended for collection, and the instrument used is free of any errors. Again, it is relevant to stress that time must be invested in identifying the best instrument or electronic system to be used, and verifying and validating this before the actual process of collecting data takes place.

Step 3: Data preparation

At this step, data is brought into the format to be evaluated. Data cleaning takes place at this stage, where the input values are checked for errors or missing values, and duplicates in data are deleted. At this step, the outliers in data are identified (this may be by using simple statistical techniques, such as averages) and these may be eliminated or smoothed out (readjusted using calculations such as regression models). Checks may also identify data that has odd values, values that may have been entered by error, and at this step this can also be corrected.

Step 4: Data analysis

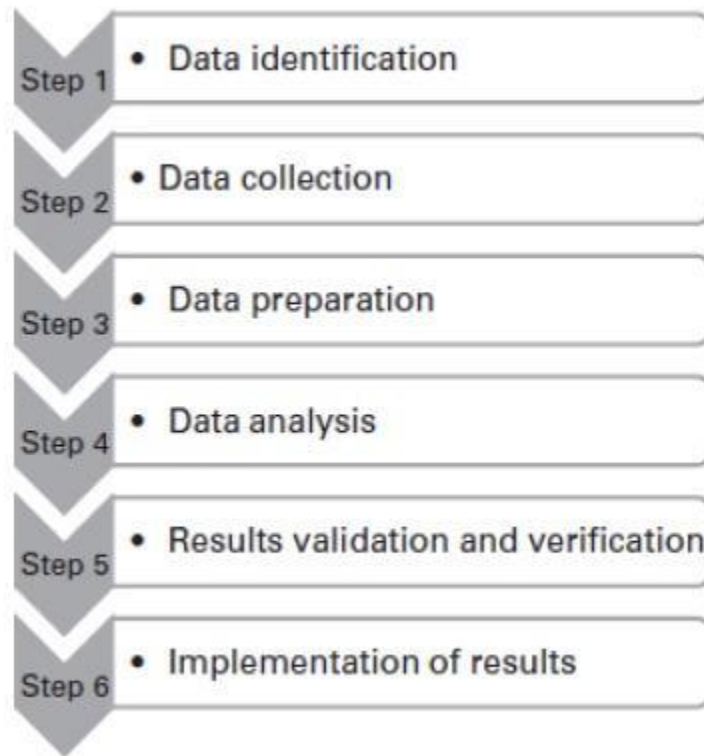
As soon as the data is clearly understood, and has been prepared in the correct format ready for analysis, the process of analyzing the data can start. Depending on the data type and format, the correct method(s) needs to be employed to analyze the data.

Step 5: Results validation and verification

After the data has been analyzed, At this stage in the process, data may present different patterns, or characteristics that will allow an analyst to construct models (for example statistical models, such as regression equations, or simultaneous equations, or particular distributions can be identified and used to generate more data, and so on).

Step 6: Implementation of results

Using models, results will be generated, therefore this step is concerned with the *results* obtained after applying particular models. The models as well as the results will need to go through a process of validation and verification of results.



Stage 1 considers data capture and management. Stage 2 is comprised of data cleaning and interrogation. These aspects are captured within the data preparation in this case. Stage 3 in their evaluation is formed of data analytics. Stage 4 considers competence set analysis. Stage 5 is built around information interpretation and decision-making.

Data type

In broad terms, data can be differentiated in two types: quantitative and qualitative.

However, we may also have mixed data that has a numerical as well as a text format.

Quantitative data has a numerical format. Quantitative data can be further split into, for example, discrete, continuous and binary data. *Discrete* data can be counted and in general has an integer value (for example: 1, 2, 3, ...). This type of data is also referred to as *categorical* data.

Continuous data is data that can be measured on a continuous scale and can have any format (for example, length, weight, temperature, and so on). This type of data is known as continuous as between any two values of this data it can be considered that there are an infinite number of other values. When evaluating processes in real time, continuous-type data can be collected. This is considered more accurate, as it gives the analyst the opportunity to observe any small variations between values, or between discrete values. However, continuous data tends to generate big volumes, and in some instances this aspect may be avoided, depending on the details required for the analysis.

Binary data in a numerical format is data that has only two values, such as 0 or 1. Binary data in the format of 0/1 is also the data that is used in computer language. One other data category that can be recognized here is *interval data* and this is measured in the form of intervals.

Quantitative data can also be expressed in *ratios*. Both ratio and interval data are continuous data types. *Qualitative data* is data that is in the format of text or a story. Qualitative methods need to be used to collect and analyse this type of data. A range of methods are used by researchers and practitioners to collect qualitative data, such as interviews, questionnaires, diary entries and others.

Data of a qualitative nature can also be observed in published case studies, summary reports, consumer feedback, and so on. A number of methods have been developed to analyse qualitative data, such as Delphi techniques, content analysis, thematic analysis and others.

Different types can be attributed to the qualitative data, such as *binary* format where only two numerical values are registered (for example yes/no, or true/false); or data can be categorized as *nominal* where different categories can be attributed to it; however, there is no specific order of rank to the data. Qualitative data is categorized as *ordinal* where an order is specified. Both nominal and ordinal data are discrete in nature.

Mixed data is data that considers numerical data, text data and any other type of data. Other type of data could be in the form of video or voice/audio, and they many need to be converted into a format that can be analysed. Different methods need to be considered in this case, where methods to analyse this type of data can be predefined, before the data has been collected, or need to be established as soon as the analysis is faced with the data received, obtained, or extracted by the system.

A number of algorithm and computer programmes are very specific on the data they can handle.

Small and big data in the supply chain

Data is the omnipresent force in all areas of the supply chain, from the movement of goods, to planning, to manufacturing, to forecasting, to the financial transaction, to services, to knowledge and information generation, to name a few of the activities. The use of reliable, accurate data has the opportunity to facilitate increased visibility in the supply chain, to allow for transparency, coordination and better communication. Supply chain analytics uses data from internal and external sources in different formats to enhance decisions, facilitate knowledge and provide up-to-date information to logistics and supply chain managers. The majority of activities and operations that need to be performed in the supply chain require data, and going further require reliable data. Data and big data are used in the supply chain to facilitate analytics such as reporting, quality control, inventory control, forecasting, production planning, distribution planning, scheduling, real-time vehicle routing, capacity planning, facility location, simulation and optimization. There are different levels of interpretation when it comes to what can be defined as small and big data. This section aims to provide further clarity to the terminology used.

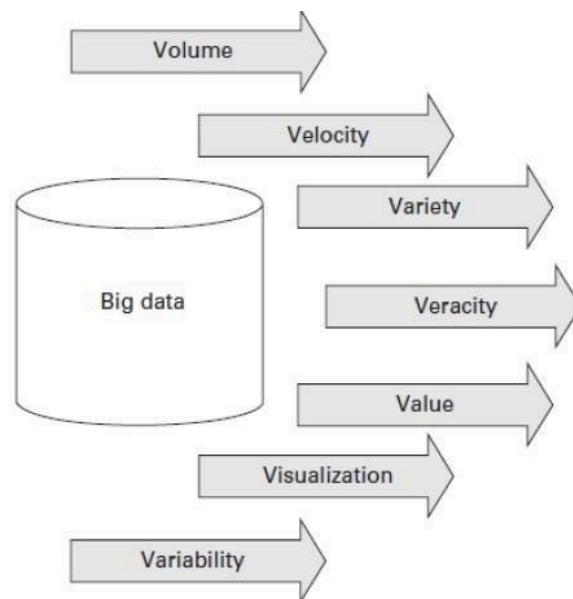
Small data

Small data tends to be required to answer a particular question, or attend to one specific goal, or be part of dedicated project (Berman, 2018). Small data in general could be owned by a research group within a particular institution, or it can be data collected for one specific project. This data may only serve the purpose of a particular aspect of a project or a particular project and will be saved in a particular file or files, following a set format with a clearly defined structure. One other characteristic of small data is that it is probably saved only for the duration of a particular project and then deleted, serving the purpose of that particular project alone. For example, in higher education, we recognize a large number of datasets with small data saved for individual students' projects, where data is collected to serve the purpose of one particular project and it is deleted after a short and specified duration of time. Data in this case may have different formats of quantitative or qualitative structure and the file formats may be of different sizes, but they could still be classed as small data sets. A large number of small datasets may collectively

result in large volumes of data, and these may be regarded by the analyst as big data due to its volume and complexity of handling. However, as the explosion of big data approaches, it is also relevant to highlight that there is incredible power in small data. There are instances where the answer to a particular question can be fully satisfied by small data without the need to make use of big data or without the requirements of big datasets. A large number of projects in academic and industrial settings are designed to only use small data to obtain the answer to the set project objectives. For a number of years, small data has governed the research arena, and it will continue to be valued and used.

Big data

Big data, as the names suggests, is the data that goes beyond the characteristics presented above of small data. Small data has a clearly defined size, structure and purpose, whereas big data is characterized by very large datasets with complex characteristics, varied structure and form, and is generated and stores in a combination of structured and unstructured formats, without necessarily a clearly defined goal.



Volume refers to the amount or size of data and databases that are available, or data that is generated and requires storage (George et al, 2016). Different social media, the internet, other devices such as

phones, sensors and so on, are in a position to generate new data in different forms and formats. New platforms have emerged that allow these kinds of data to be stored and manipulated.

Velocity represents the speed at which the data is being generated and made available. A number of devices are set to generate data and this is being done at an incredible speed (George et al, 2016). This may be from different set devices such as controllers, to agents that are set to indicate when, for example, inventory has reached a certain level, simulations are also set to generate data and many others. However, a number of questions still remain: Will all generated data be saved? Or used? Will it have the potential to generate value?

Variety is the type of data that is being created in structured, semi-structured or unstructured format. Data in varied forms has been used over a number of years, and there is an immense volume of varied types of data. Within a supply chain, data is being collected from a number of different sources (such as point of sale data, membership cards, RFID tags, GPS locations, barcodes, etc), each coming with a different data format. At the same time, organizations recognize more and more the value hidden in unstructured data, where new analytical techniques are being adopted to provide analysis of data of unstructured format.

Veracity refers to the way in which the correctness, accuracy and quality of data is verified. Taking this step confirms the reliability of the data (Ittmann, 2015).

Value is the added worth or usefulness data can generate as a result of the information provided. The availability of data alone is not sufficient if this can't be turned into value.

Visualization refers to the way in which data is represented, to make better use of what the data represents, to enhance the decision-making process. Easy accessibility and readability of data allows for better decision-making.

Variability in data indicates that data's meaning may change due to analysis or over a short period of time. To understand variability in data, Sivarajah et al (2017) mention the use of sentiment analysis

Challenges in working with data

1. Data Collection Challenges

- **Data availability:** Relevant data might not exist or may be hard to access.
- **Data silos:** Data stored in separate systems or departments can be hard to integrate.
- **Inconsistent formats:** Different sources may use different formats or standards.
- **Ethical and legal concerns:** Data privacy laws (e.g., GDPR) can limit data access and usage.

2. Data Quality Issues

- **Missing values:** Important information may be absent or incomplete.
- **Noise and errors:** Data can contain typos, outliers, or inaccurate entries.
- **Duplicates:** Repeated entries can skew results.
- **Outdated information:** Data may be stale, especially in fast-changing environments.

3. Data Processing Difficulties

- **Data cleaning complexity:** Time-consuming and sometimes requires domain knowledge.
- **Data integration:** Merging data from different sources can introduce inconsistencies.
- **Handling large volumes:** Big data can be difficult to process and store efficiently.

4. Technical and Infrastructure Barriers

- **Lack of tools or technology:** Not all organizations have the resources for advanced analytics.
- **Scalability issues:** Systems may not handle growth in data size or user demand.
- **Security risks:** Storing and transmitting data safely is a constant challenge.

5. Analytical and Modeling Issues

- **Bias in data or models:** Can lead to unfair or incorrect results.
- **Choosing the right methods:** Selecting appropriate algorithms or statistical techniques is critical.
- **Interpretability:** Complex models (e.g., deep learning) are often hard to explain.

6. Organizational and Human Factors

- **Lack of data literacy:** Stakeholders may struggle to understand or use data effectively.
- **Resistance to change:** People may distrust data-driven decisions or prefer traditional methods.
- **Miscommunication:** Data findings may be misinterpreted or poorly communicated.

7. Ethical and Social Challenges

- **Privacy concerns:** Especially when dealing with personal or sensitive data.
- **Algorithmic fairness:** Ensuring that models do not discriminate.
- **Transparency:** Making data and algorithms understandable and accountable.

