UE 4 - Logistique et Gestion de la Supply Chain
UE 4 - Logistics and Supply Chain Management

Module 4.3 - Modélisation & optimisation de la Supply Chain
Module 4.3 - Supply chain modelling and optimization

Riccardo Rossi
Massimiliano Gastaldi
Riccardo Ceccato, riccardo.ceccato@dicea.unipd.it

Exam

1. Multiple choice questions:
   - Choose one answer among 3 or 4 alternatives
   - Correct answer: 1 point
   - Wrong answer or no answer: 0 point
2. Exercises:
   - Needed formulas are provided
   - Calculate the result and report the calculation
   - Choose the correct answer among 3 or 4 alternatives
   - Correct answer: 2 points
   - Wrong answer or no answer: 0 point

The text of the exam will be provided in English and French.

Exam

1. Example of multiple choice questions:

   What is the correct order of the three main phases of decision making in supply chain:

   - [x] (1) Supply chain strategy or design, (2) supply chain planning, (3) supply chain operations
   - [ ] (1) Supply chain planning, (2) supply chain operations, (3) supply chain strategy or design
   - [ ] (1) Supply chain operations, (2) supply chain planning, (3) supply chain strategy or design

2. Examples of exercises:

   A truck costs 500$ per day independently from the actual loaded weight and it carries at most 25 tons. What is the total cost if you need to carry 35 tons?

   - [ ] 500$
   - [x] 1000$
   - [ ] 2000$

   Write here your calculation:
   35/25 = 1.4; 1.4>1 -> 2 trucks are needed
   2 x 500$ = 1000 $

   A company sold 5 units (1 year ago), 4 units (2 years ago), 6 units (3 years ago) and 7 units (4 years ago). Forecast the sold units for the current year by using Simple Moving Average. The formula is:

$$F_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \ldots + A_{t-n}}{n}$$

F_t = Forecast for the coming period
n = Number of periods to be averaged
A_{t-1} = Actual occurrence in the past period
A_{t-2}, A_{t-3}, and A_{t-n} = Actual occurrences two periods ago, three periods ago, and so on, up to n periods ago

   - [ ] 3.5 units
   - [x] 5.5 units
   - [ ] 10.5 units

   Write here your calculation:

$$\frac{5 + 4 + 6 + 7}{4} = 5.5 \text{ units}$$

<u>Course structure</u>

Theoretical part:

- Introduction
- Distribution and transportation networks design
- Facility location planning
- Capacity management
- Sourcing strategy
- Production strategy
- Demand forecasting
- Inventory management
- Risk management
- Optimization of supply chain

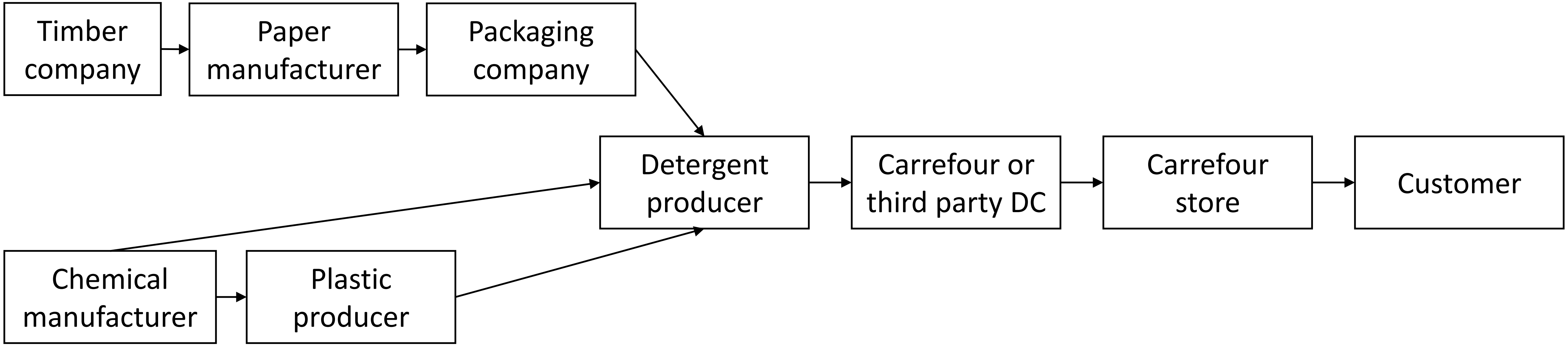
Practical part:

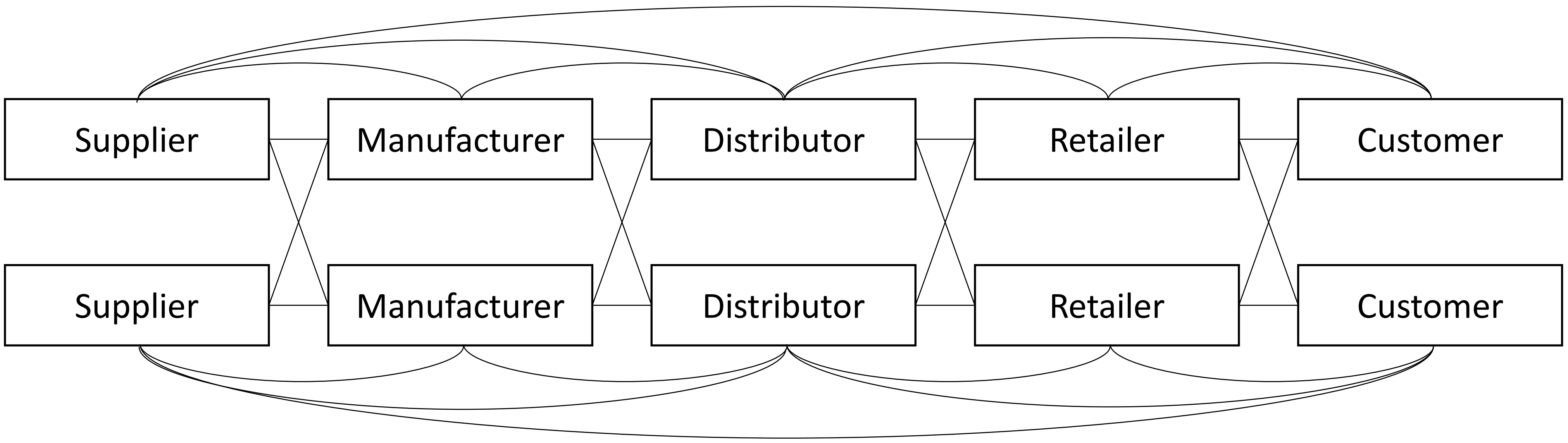
- Exercises
- Software to model and optimize supply chain:
  - Facility location problem
  - Network optimization
  - Supply chain simulation (supply chain performance)
  - Risk modelling

Definition of a supply chain
A supply chain consists of all parties involved, directly or indirectly, in fulfilling a customer request and it includes all functions involved in receiving and filling that request (e.g. new product development, marketing, operations, distribution, finance, and customer service).

A supply chain involves the constant flow of information, product, and funds among different stages.

The customer is an integral part of the supply chain.





A typical supply chain may involve the following stages:

- Customers
- Retailers
- Wholesalers/distributors
- Manufacturers
- Component/raw material suppliers

Each stage in a supply chain is connected through the bidirectional flow of products, information, and funds.

Objective of a supply chain

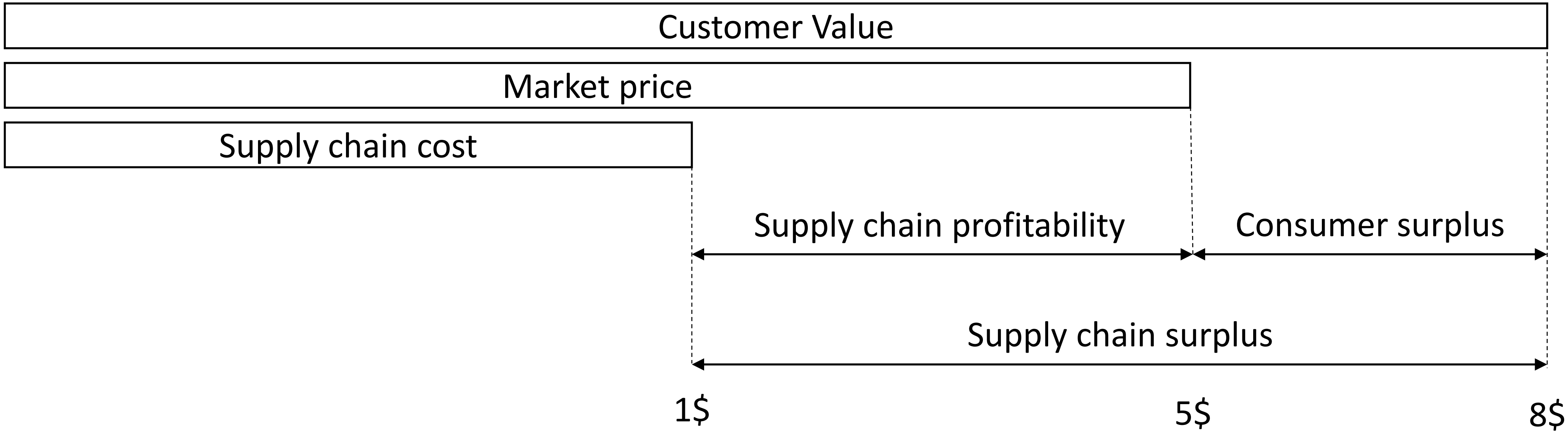
The objective of every supply chain should be to maximize the overall value generated (Supply chain surplus).

$$Supply\ Chain\ Surplus\ =\ Customer\ Value\ -\ Supply\ Chain\ Cost$$

The value of the final product:

- may vary for each customer
- can be estimated by the maximum amount the customer is willing to pay for it

$$Customer\ Value\ -Market\ price = Consumer\ surplus$$



Supply chain success should be measured in terms of supply chain surplus and not in terms of the profits at an individual stage.

The customer is the only source of revenue.

All other cash flows are simply fund exchanges that occur within the supply chain, given that different stages have different owners.

Effective supply chain management involves the management of supply chain assets and product, information, and fund flows to grow the total supply chain surplus.

Decisions in a supply chain

Decisions fall into three phases, depending on the frequency of each decision and the time frame during which a decision phase has an impact.

1. Supply chain strategy or design: how to structure the supply chain over the next several years:

- what the chain’s configuration will be
- how resources will be allocated
- what processes each stage will perform

A firm must ensure that the supply chain configuration supports its strategic objectives and increases the supply chain surplus during this phase.

Supply chain design decisions are typically made for the long term (a matter of years) and are expensive to alter on short notice.

Uncertainty in anticipated market conditions over the following few years must be taken into account.

2. <u>Supply chain planning</u>: the time frame considered is a quarter to a year. The supply chain's configuration determined in the strategic phase is fixed.

This configuration establishes constraints within which planning must be done.

The goal of planning is to maximize the supply chain surplus that can be generated over the planning horizon given the constraints established during the strategic or design phase.

Uncertainty in demand, exchange rates, and competition over this time horizon must be considered.

The result is a set of operating policies that govern short-term operations.

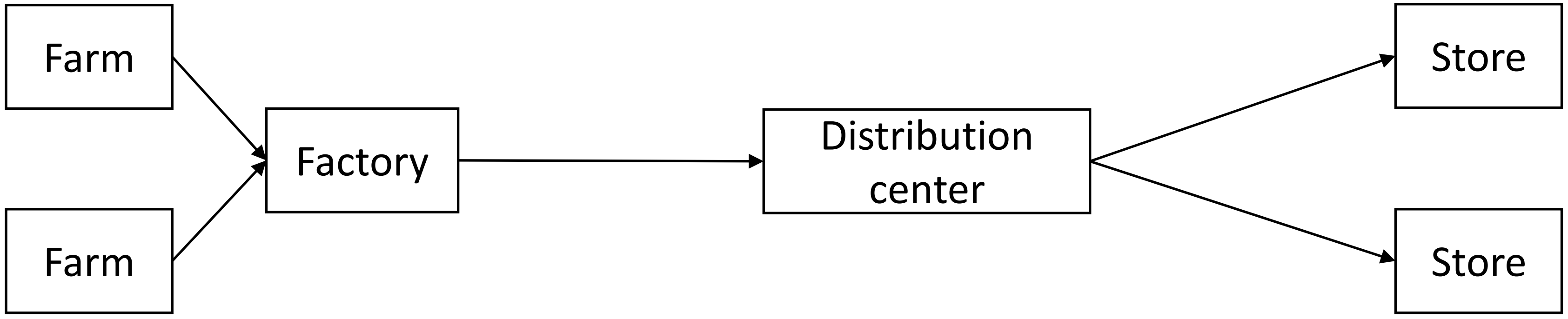
3. <u>Supply chain operation</u>: the time horizon is weekly or daily. Companies make decisions regarding individual customer orders.

Supply chain configuration is considered fixed and planning policies are already defined.

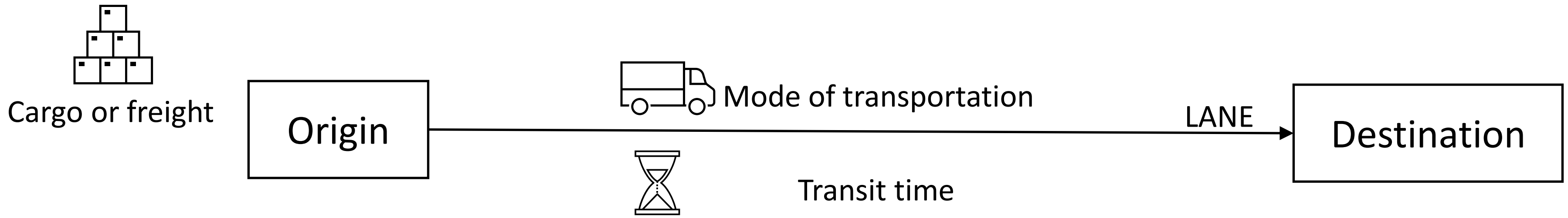
There is less uncertainty about demand information.

The goal of supply chain operations is to handle incoming customer orders in the best possible manner.

Transportation refers to the movement of product from one location to another as it makes its way from the beginning of a supply chain to the customer.



Transportation is a significant component of the costs incurred by most supply chains (2-3% of the total revenues and 6–8% of the company’s costs).



Main actors

- Shipper:
  - It requires the movement of the product
  - It tries to maximize the return from its assets (e.g. locomotives, trucks, airplanes)
- Carrier:
  - It moves or transports the product
  - It tries to minimize the total cost while providing an appropriate level of responsiveness to the customer
- Owner of transport infrastructures
- Transport policy maker:
  - It addresses national resources
  - It prevents abuse of monopoly power
  - It promotes fair competition
  - It balances environmental, energy, and social concerns

Factors affecting network design decisions

Strategic factors

- Focus on cost leadership: finding the lowest cost location for manufacturing facilities
- Convenience store: network with small many stores that cover an area
- Discount store: network with large few stores that cover an area
- Global supply chain: facilities in different countries playing different roles

Technological factors

- If production technology displays significant economies of scale, a few high-capacity locations are most effective
- If facilities have lower fixed costs, many local facilities are preferred because this helps lower transportation costs

Macroeconomic factors

As global trade has increased, macroeconomic factors have had a significant influence on the success or failure of supply chain networks.

- Tariffs and tax incentives. Tariffs refer to any duties that must be paid when products and/or equipment are moved across international, state, or city boundaries. High tariffs lead to more production locations within a supply chain network, with each location having a lower allocated capacity. Tax incentives are a reduction in tariffs or taxes that countries, states, and cities often provide to encourage firms to locate their facilities in specific areas. Developing countries often create free trade zones in which duties and tariffs are relaxed as long as production is used primarily for export; they also provide additional tax incentives based on training, meals, transportation, and other facilities offered to the workforce. Many countries also place minimum requirements on local content and limits on imports to help develop local manufacturers.
- Exchange-rate and demand risk. Fluctuations in exchange rates are common. To take advantage of exchange-rate fluctuations and increase profits, an effective way is to build some overcapacity into the network and make the capacity flexible so it can be used to supply different markets. Companies must also take into account fluctuations in demand caused by changes in the economies of different countries.
- Freight and fuel costs. High fluctuations are best dealt with by hedging prices on commodity markets or signing suitable long-term contracts.

Political Factors

- Companies prefer to locate facilities in politically stable countries where the rules of commerce and ownership are well defined
- Global Political Risk Index: it aims to measure the capacity of a country to withstand shocks or crises along government, society, security, and economy

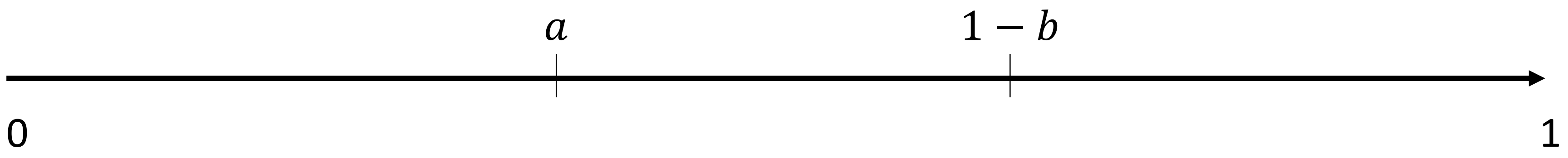
Infrastructure Factors

- Availability of sites and labor, proximity to transportation terminals, rail service, proximity to airports and seaports, highway access, congestion, and local utilities

Competitive Factors

A fundamental decision firms make is whether to locate their facilities close to or far from competitors. The form of competition and factors such as raw material or labor availability influence this decision.

- Positive externalities. Positive externalities occur when the collocation of multiple firms benefits all of them. Positive externalities lead to competitors locating close to each other.
- Locating to split the market. When there are no positive externalities, firms locate to be able to capture the largest possible share of the market. Simple model (Hotelling):



$$d_1 = a + \frac{1 - b - a}{2} \qquad d_2 = \frac{1 + b - a}{2}$$

- Both firms maximize their market share if they move closer to each other and locate at a = b = 1 /2.
- The result of competition is for both firms to locate close together, even though doing so increases the average distance to the customer.
- If the firms compete on price and the customer incurs the transportation cost, it may be optimal for the two firms to locate as far apart as possible.

Customer Response Time and Local Presence

- Firms that target customers who value a short response time must locate close to them.
- If a firm is delivering its product to customers, use of a rapid means of transportation allows it to build fewer facilities and still provide a short response time. This option, however, increases transportation cost.

Logistics and Facility Costs

Logistics and facility costs incurred within a supply chain change as the number of facilities, their location, and capacity allocation change.

- Inventory and facility costs increase as the number of facilities in a supply chain increases. Transportation costs decrease as the number of facilities increases. If the number of facilities increases to the point at which inbound economies of scale are lost, then transportation costs increase.
- The supply chain network design is also influenced by the transformation occurring at each facility. When there is a significant reduction in material weight or volume as a result of processing, it may be better to locate facilities closer to the supply source rather than the customer.
- The facilities in a supply chain network should at least equal the number that minimizes total logistics cost. A firm may increase the number of facilities beyond this point to improve the response time to its customers.

Introductory case study

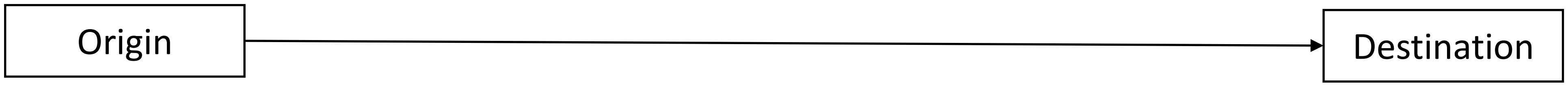
[From Ivanov et al. (2019)]

- A large amount of the wood is bought from Asian suppliers.
- 3 container ports
- The container is transshipped on hired trucks, which bring it to the Wood Company manufacturing site close to Munich
- Transportation costs for the raw materials are quite high. Wood Company does not possess its own trucks. Whenever a certain transport quantity has to be moved, Wood Company hires a truck from a freight forwarding company at a fixed daily rate. Is this a good idea?
- Wood Company are preparing the integration of two other furniture factories into their company. To ensure the availability of sufficient quantities of wood, imports will also be processed via the Dutch port of Rotterdam.
- Should the previous method of transport be retained? What are the possible modifications required for hinterland transportation? Should Wood Company buy its own trucks and set up its own transport department? Which production site should be supplied from which port in order to keep transportation costs to a minimum? Should all necessary transport services be given away to a transport service provider?

Generic Transport Network Structures

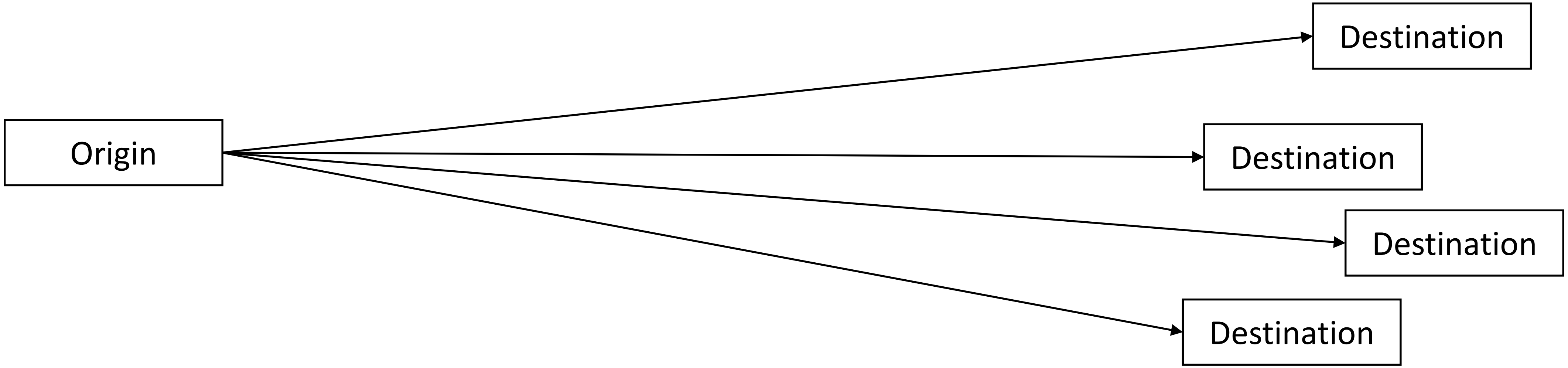
<u>Transportation link</u>

- The supply quantities are available at a single location (the origin) and the available supply quantity is requested completely or partially from a single location (the destination)
- There is no opportunity to optimize (reduce) the transportation efforts if the destination wants to receive the requested quantity immediately
- Time of delivery of individual shipments can be arranged



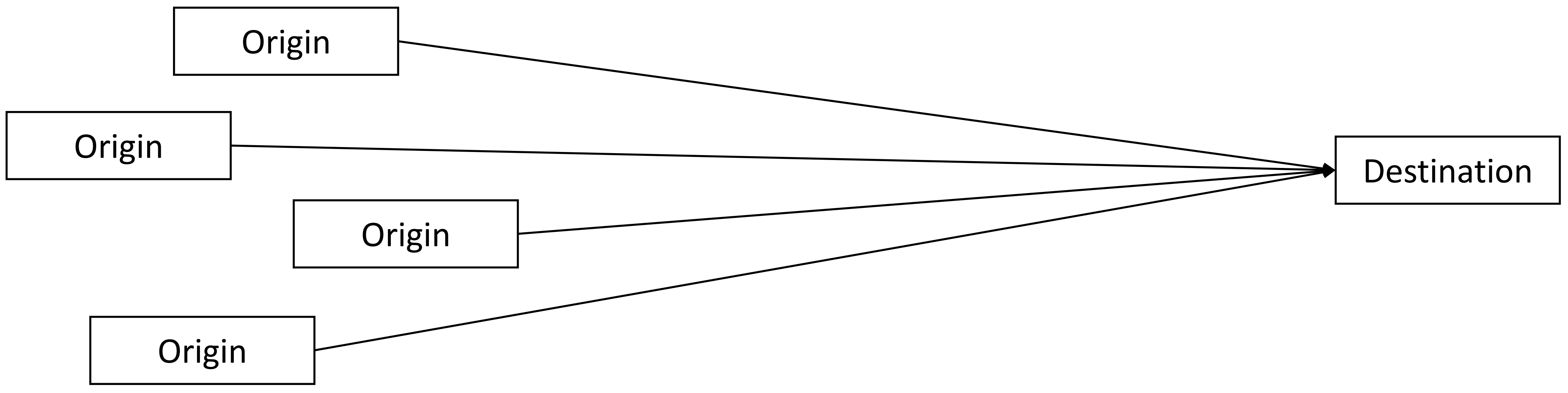
Distribution network

- Decisions about the distribution network are the same as in the transportation link
- Only the time of delivery can be decided



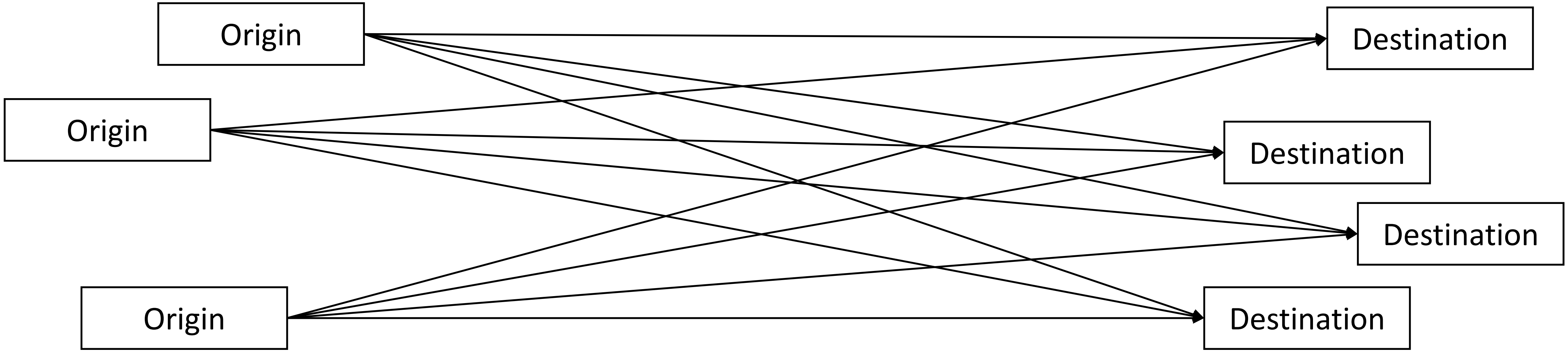
Supply network

- The destination requests a certain supply quantity, but the demand quantity is distributed over several origin locations
- It is necessary to decide which origin should provide which supply quantity



Many-to-many network

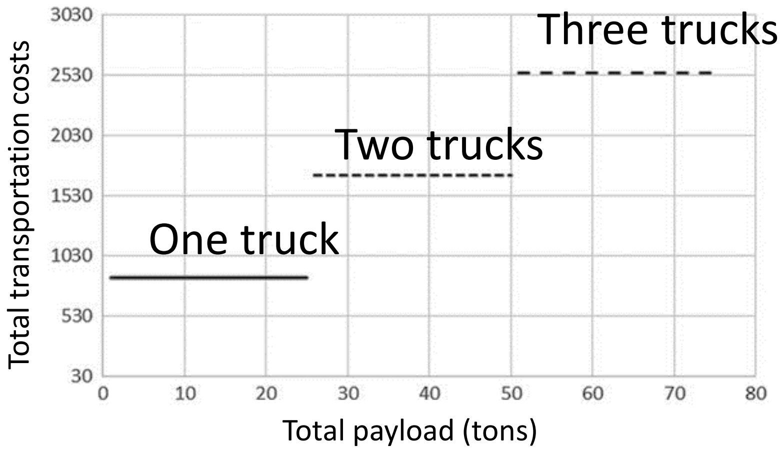
- Several origins provide supply quantities, but there are also several demanding destination locations
- It is necessary to decide from which location(s) an individual destination location is supplied
- For each origin location there is a decision required about the quantities and destinations that are supplied



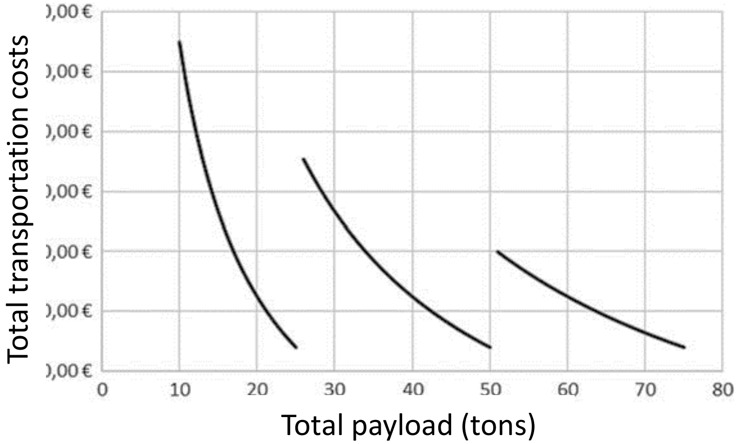
Economies of scale in transportation

Economies of scale are typically realized if large quantities of a product are bought or if unnecessary setups of machines can be avoided.

- A truck costs 850 € per day (independently from the actual loaded weight)
- It provides a payload of at most 25 tons



- There are discontinuities of the total costs whenever another truck is needed



- If the costs for the truck are equally split over the complete payload assigned to one truck, then the average costs per ton decrease with increasing payload
- If the payload exceeds 25 tons it becomes necessary to hire an additional truck which causes a discontinuity in the total costs leading to an implied increase in the costs per ton payload

Consolidation of shipments

Consolidation is the combination of two or even more quantities to be shipped between a given origin and a given destination.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	32	53	18	34	43	60	30	270
Fulfilled demand	32	53	18	34	43	60	30	270
Hired trucks	2	3	1	2	2	3	2	15
Costs (€)	1700	2550	850	1700	1700	2550	1700	12,750
Costs (per ton) (€)	53.13	48.11	47.22	50.00	39.53	42.50	56.67	

- 15 trucks to be hired
- Weekly transportation costs is 12,750 € at an average fill rate of 71.5%

Postponement

The incoming demand is merged with postponed quantities to the recent available demand. Full truck loads (FTL) are derived from the available demand quantities, but residual quantities are postponed until a later day of the week. This postponement strategy is fragile.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
Incoming OD-demand	32	53	18	34	43	60	30	270
Postponed demand	0	7	10	3	12	5	15	
Available demand	32	60	28	37	55	65	45	
Now fulfilled demand	25	50	25	25	50	50	45	270
Hired trucks	1	2	1	1	2	2	2	11
Costs (€)	850	1700	850	850	1700	1700	1700	9350
Costs (per ton) (€)	34.00	34.00	34.00	34.00	34.00	34.00	37.78	

- 11 trucks to be hired
- Weekly transportation costs is 9350 € at an average fill rate greater than 98%

Demand waiting in Hamburg

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	32	53	18	34	43	60	30	270
Fulfilled demand	32	53	18	34	43	60	30	270
Hired trucks	2	3	1	2	2	3	2	15
Costs (€)	1700	2550	850	1700	1700	2550	1700	12,750
Costs (per ton) (€)	53.13	48.11	47.22	50.00	39.53	42.50	56.67	48.17
Utilization (%)	64	71	72	68	86	80	60	72

Demand waiting in Wilhelmshaven

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	13	12	25	13	26	38	11	138
Fulfilled demand	13	12	25	13	26	38	11	138
Hired trucks	1	1	1	1	2	2	1	9
Costs (€)	850	850	850	850	1700	1700	850	7650
Costs (per ton) (€)	65.38	70.83	34.00	65.38	65.38	44.74	77.27	60.43
Utilization (%)	52	48	100	52	52	76	44	61

Demand waiting in Bremerhaven

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	23	42	31	28	38	50	20	232
Fulfilled demand	23	42	31	28	38	50	20	232
Hired trucks	1	2	2	2	2	2	1	12
Costs (€)	850	1700	1700	1700	1700	1700	850	10,200
Costs (per ton) (€)	36.96	40.48	54.84	60.71	44.74	34.00	42.50	44.89
Utilization (%)	92	84	62	56	76	100	80	79

- 36 trucks to be hired
- Weekly transportation costs is 30,600 € at an average fill rate between 61% and 79%
- Average costs per moved ton vary between 44.89 € and 60.43 €
- Only 17 of the 36 hired trucks are fully loaded

Milk-run

One truck visits two or more pickup locations before bringing the whole load to a common destination. Milk-runs possible:

(1) Hamburg, Bremerhaven and Wilhelmshaven are served as pickup locations by the same truck
(2) Hamburg and Bremerhaven are served in one milk-run
(3) Demand quantities waiting in Hamburg and in Wilhelmshaven are consolidated
(4) One truck loads demand quantities in Bremerhaven and in Wilhelmshaven

Demand waiting in Hamburg

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	32	53	18	34	43	60	30	270
Fulfilled demand	25	50	0	25	25	50	25	200
Hired trucks	1	2	0	1	1	2	1	8
Costs (€)	850	1700	–	850	850	1700	850	6.800
Utilization (%)	100	100	0	100	100	100	100	86

Demand waiting in Bremerhaven

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	23	42	31	28	38	50	20	232
Fulfilled demand	0	25	25	25	25	50	0	150
Hired trucks	0	1	1	1	1	2	0	6
Costs (€)	–	850	850	850	850	1700	–	5100
Utilization (%)	0	100	100	100	100	100	0	71

Demand waiting in Wilhelmshaven

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	13	12	25	13	26	38	11	138
Fulfilled demand	0	7	25	0	25	25	11	93
Hired trucks	0	1	1	0	1	1	1	5
Costs (€)	–	850	850	–	850	850	850	4250
Utilization (%)	0	28	100	0	100	100	44	53

Remaining demand fulfilled in milkruns

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
Milkrun demand in Hamburg (HH)	7	3	18	9	18	10	5	70
Milkrun demand in Bremerhaven (BHV)	23	17	6	3	13	0	20	82
Milkrun demand in Wilhelmshaven (WHV)	13	5	0	13	1	13	0	45
Total demand	43	25	24	25	32	23	25	
Milkrun HH-BHV-WHV	25	25		25				
Milkrun HH-BHV			24		25		25	
Milkrun HH-WHV						23		
Milkrun BHV-WHV	18				7			
Fulfilled demand	43	25	24	25	32	23	25	
Hired trucks	2	1	1	1	2	1	1	9
Costs (€)	1700	850	850	850	1700	850	850	7650
Utilization (%)	86	100	96	100	64	92	100	91
Overall number of hired trucks	3	5	3	3	5	6	3	28
Costs (€)	2550	4250	2550	2550	4250	5100	2550	23,800
Moved quantity	68	107	74	75	107	148	61	640
Costs per ton (€)	38	40	34	34	40	34	42	37.38
Utilization (%)	91	86	99	100	86	99	81	92

- 28 trucks to be hired
- Weekly transportation costs 23,800 € at an average fill rate of 92%
- Average costs per ton of 37.38 €

However:

- From day to day, the milk-runs have to be adapted to recent demand
- If the total quantities from the milk-run locations are slightly above the full truck load quantity, then the incorporation of an almost empty truck cannot be avoided

<u>Transshipment</u>
Goods are transferred from one truck to another one, to increase the fill rate of trucks travelling long distances.

- The trucks travelling from the ports to the transshipment facility travel relatively short distances, therefore, they can shuttle two or even three times between their pickup port and Hanover, so a single truck is enough to forward all quantities from the port to the transshipment facility in the port hinterland.
- Due to the consolidation of the inbound flow of quantities originating from the three sea ports, there are huge opportunities to compile full truck loads for the main-haul service between Hanover and Munich

Possible set-ups for a hired truck (available and paid for 24 h)

A. Three feeder cycles between a port and the transshipment facility and a subsequent main-haul trip from Hanover to Munich
B. Six feeder cycles and no main-haul trip
C. One main-haul trip only

Reduced tariffs:

- 750 €, since the truck operator incurs fewer costs for overnight stays for the vehicle and crew (Type B pattern)
- 700 €, to a truck executing one main haul only due to reduced travel distances (Type C pattern)

Demand waiting in Hamburg

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	32	53	18	34	43	60	30	270
Fulfilled demand	32	53	18	34	43	60	30	270
Required feeder services	2	3	1	2	2	3	2	15

Demand waiting in Bremerhaven

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	23	42	31	28	38	50	20	232
Fulfilled demand	23	42	31	28	38	50	20	232
Required feeder services	1	2	2	2	2	2	1	12

Demand waiting in Wilhelmshaven

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
OD-demand (in tons)	13	12	25	13	26	38	11	138
Fulfilled demand	13	12	25	13	26	38	11	138
Required feeder services	1	1	1	1	2	2	1	9

$$Z = 850y_A + 750y_B + 700y_C$$

$$3y_A + 6y_B \geq k_{feed}$$

$$1y_A + 1y_C \geq k^{mh}$$

$$y_A, y_B, y_C \in \{0; 1; 2; \dots\}$$

	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Sums
Total demand Hanover to Munich	68	107	74	75	107	148	61	640
Required feeder cycles	4	6	4	5	6	7	4	36
Required feeder cycles	3	5	3	3	5	6	3	28
Hired truckes for type-A-pattern	2	5	2	2	2	3	2	18
Hired truckes for type-B-pattern	0	0	0	0	0	0	0	0
Hired truckes for type-C-pattern	1	0	1	1	3	3	1	10
Available feeder cycles	6	15	6	6	6	9	6	54
Available main-haul services	3	5	3	3	5	6	3	28
Hiring costs (€)	2400	4250	2400	2400	3800	4650	2400	22,300
Hired trucks	3	5	3	3	5	6	3	28

- 28 trucks to be hired
- Weekly transportation costs is reduced by 6% compared to milk-runs

However:

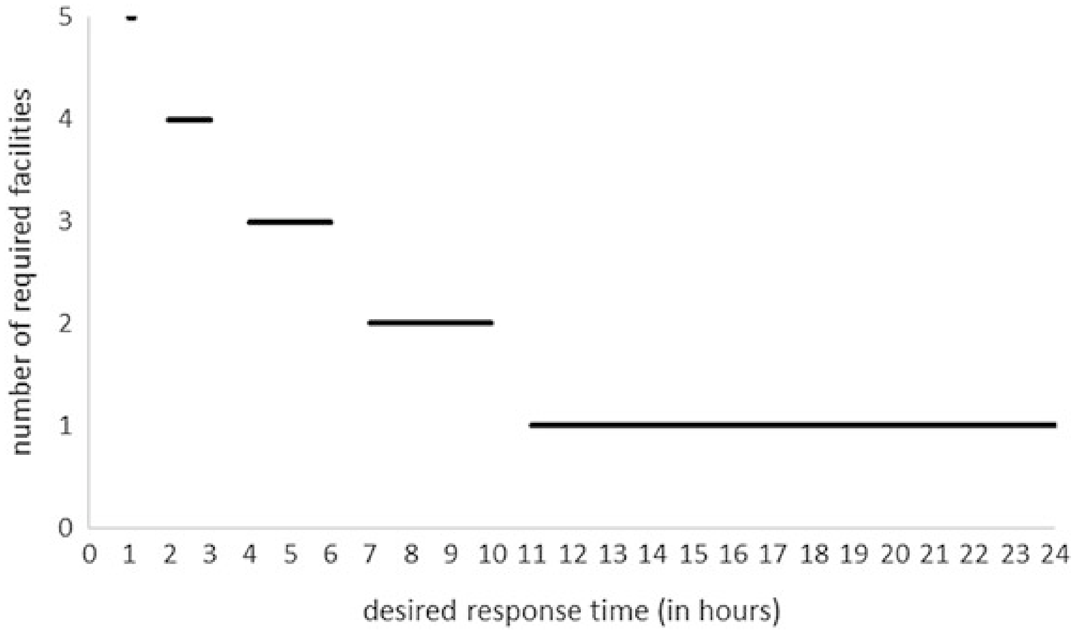
- The transshipment facility must be constructed, hired, or rented
- The transshipment activities and the necessary storage processes incur costs

General directions for designing and evaluating transportation networks:

1. customer needs must be identified, but they are unknown, hidden, or will change
2. the costs of the installation and deployment of the transportation network have to be determined, but they are subject to external impacts, from oscillating customer demand, varying fuel prices, labor costs, …

Therefore, only a rough estimation of benefits with regard to customer satisfaction and costs is possible.

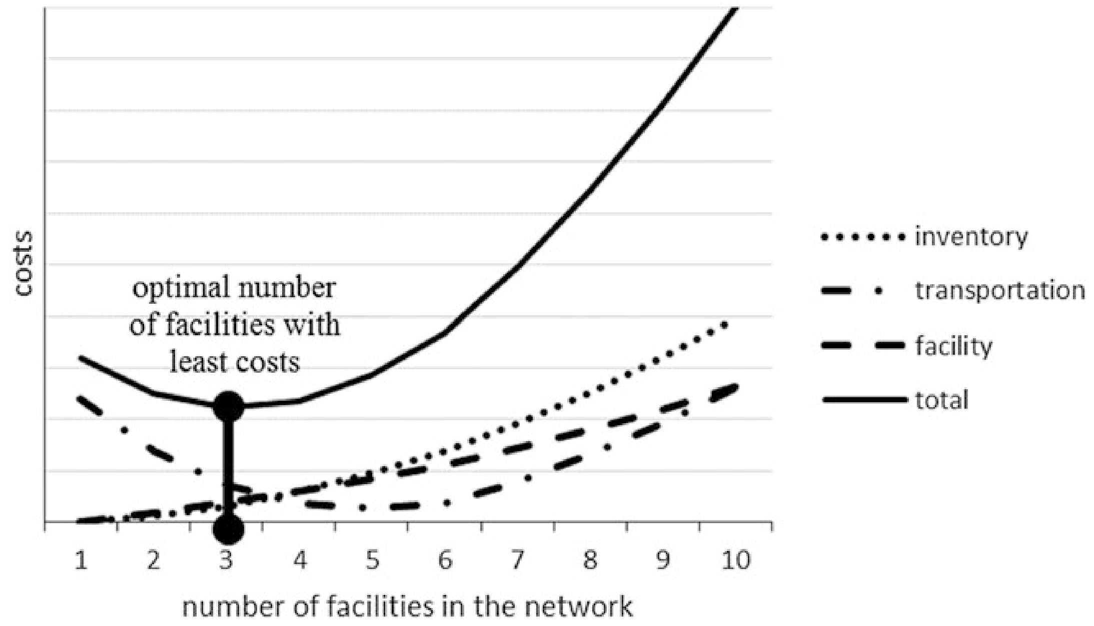
The major driver of customer satisfaction is the ability to react immediately and with high reliability to an expressed customer demand.



- The average re-supply time or response time, which equals the difference between the time of arrival of the additional material and the time when the need is expressed, tends to decrease with the increasing number of network nodes
- There is a strong relationship between the number of facilities for storing supply materials and the customer service level
- Costs for keeping all facilities running increases if the number of operated facilities is increased

There are three major cost drivers to be considered during the determination of the right network design:

1. Inventory costs: they tend to grow if the number of networked facilities is increased because each facility holds a certain quantity of inventory
2. Facility costs: they increase if the number of facilities is increased
3. Transportation costs:
   - up to a certain number of maintained transshipment facilities, additional economies of scale can be realized by consolidation
   - beyond a threshold value, it is no longer possible to completely fill a truck. The average quantity to be moved by a truck decreases, since total transport demand is distributed over a higher number of transport options, so that the costs per moved ton increase further



The function for the total costs is convex, with an optimal number of facilities with minimal total costs. Incorporating this optimal number of facilities minimizes the total costs of the transportation system and thus contributes to balancing costs and customer satisfaction.

The performance of a distribution network should be evaluated along two dimensions:

1. Value provided to the customer
2. Cost of meeting customer needs

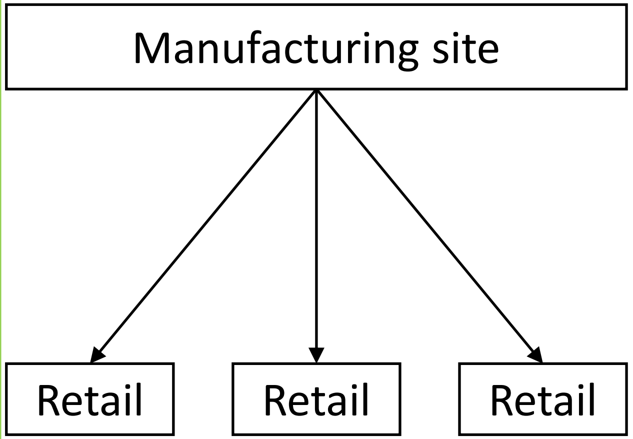
The customer needs that are met influence the company's revenues, which, along with cost, decide the profitability of the delivery network.

Measures that are influenced by the structure of the distribution network:

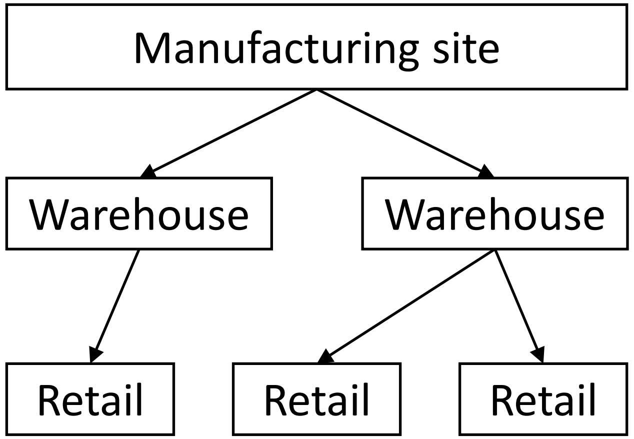
- Response time: the amount of time it takes for a customer to receive an order
- Product variety: the number of different products or configurations that are offered by the distribution network
- Product availability: the probability of having a product in stock when a customer order arrives
- Customer experience: the ease with which customers can place and receive orders and the extent to which this experience is customized
- Time to market: the time it takes to bring a new product to the market
- Order visibility: the ability of customers to track their orders from placement to delivery
- Returnability: the ease with which a customer can return unsatisfactory merchandise and the ability of the network to handle such returns

Main types of distribution networks

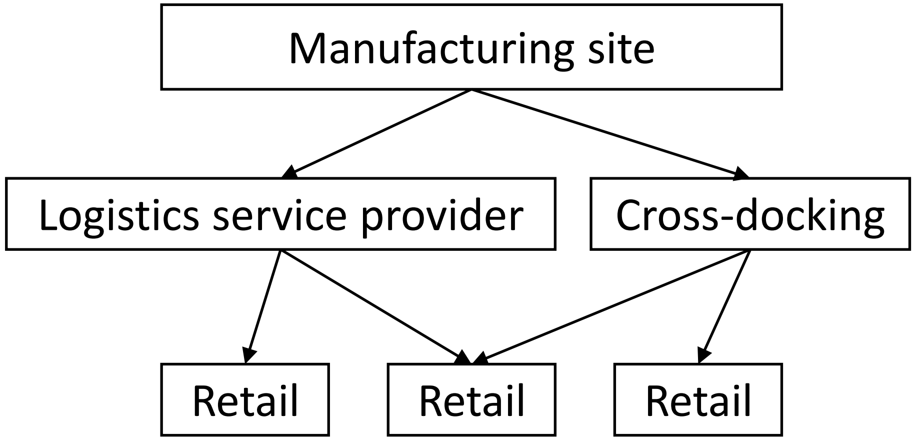
- Direct network with manufacturer storage
- Multi-stage network with distributor storage
- Multi-stage network with cross- docking and logistics service provider



Direct network with manufacturer storage



Multi-stage network with distributor storage



Multi-stage network with cross- docking and logistics service provider

Cross-docking

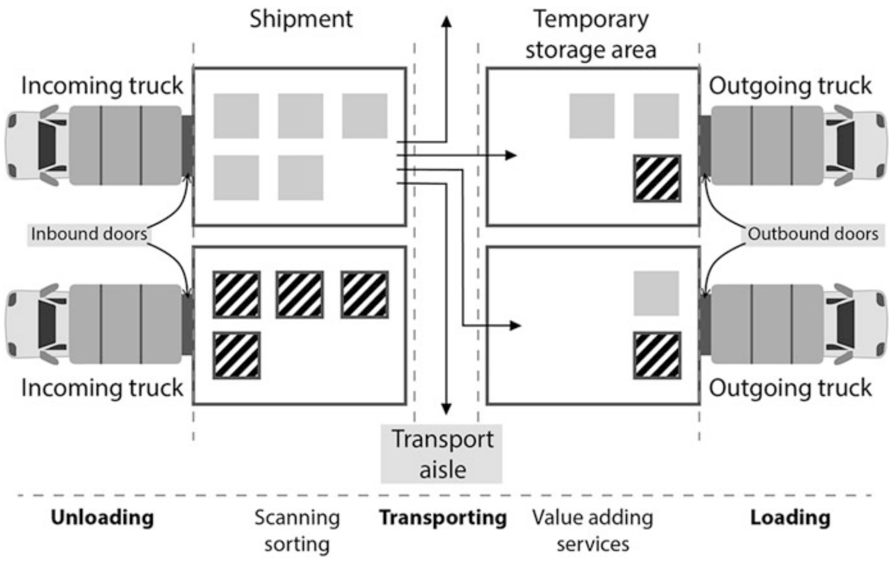
It implements the idea of the consolidation of goods from inbound trucks to outbound trucks via an intermediate transit point. Goods do not remain in the warehouse longer than 24 h.

Advantages:

The realization of economies of scale in transportation, frequent bundled deliveries, inventory reduction, faster product flow, reduction of errors, savings in material handling and labor costs, increase in fleet capacity utilization.

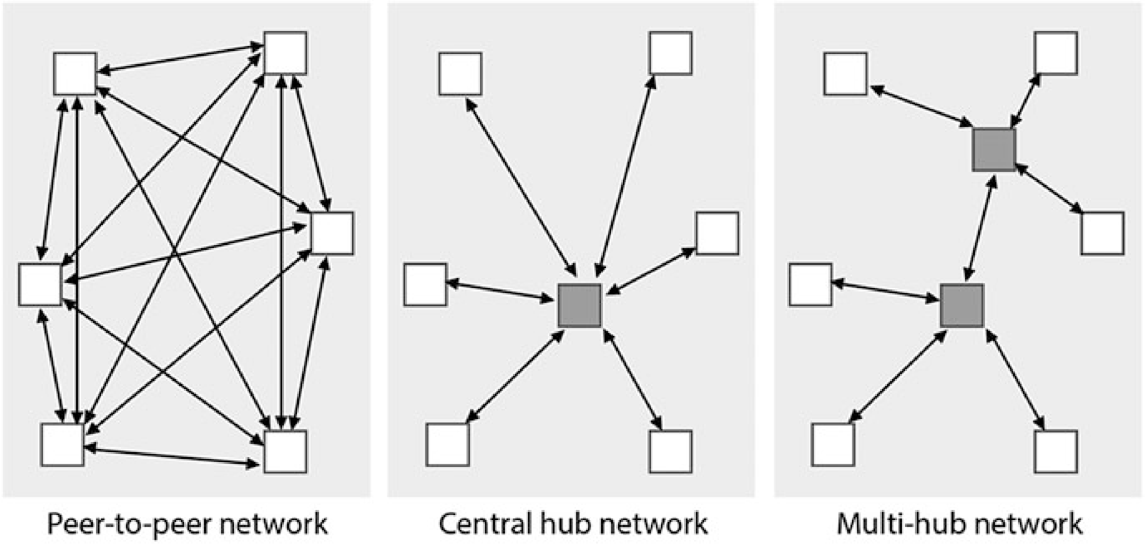
Disadvantages:

- A high investment is required.
- The coordination complexity and risks of stock-out/disruptions in the supply chain also increase.
- Data security issues become crucial.



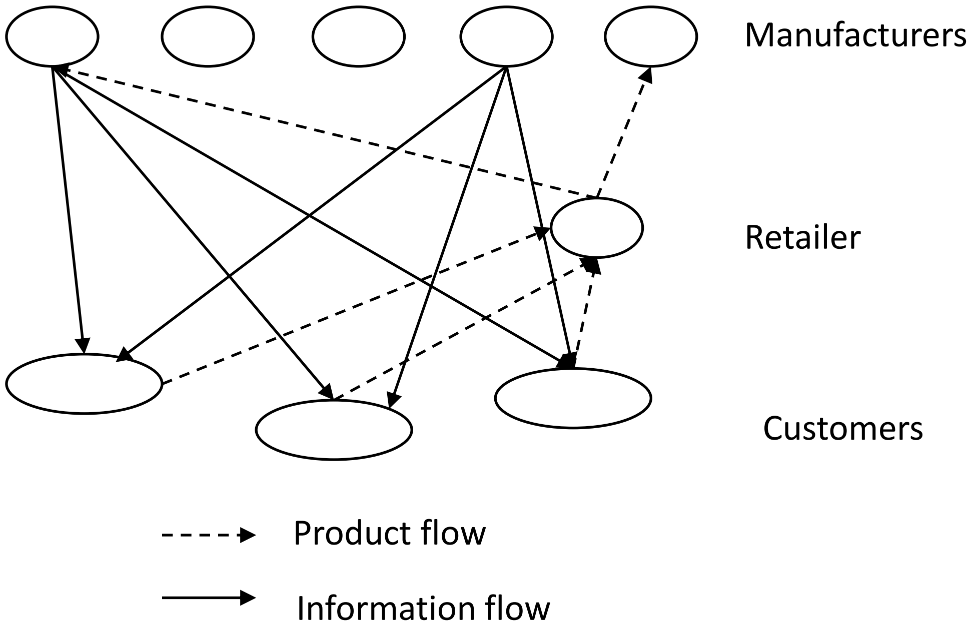
Hub-and-spoke

It also implements the idea of the consolidation. A hub represents the consolidation element in the distribution network. Spokes are the regional warehouses and customers. This concept is used for shipment consolidation with similar advantages and disadvantages as cross-docking.



Detailed types of distribution networks

- Manufacturer storage with direct shipping
- Manufacturer storage with direct shipping and in-transit merge
- Distributor storage with carrier delivery
- Distributor storage with last-mile delivery
- Manufacturer/distributor storage with customer pickup
- Retail storage with customer pickup



<u>Manufacturer Storage with Direct Shipping (drop-shipping)</u>

- Product is shipped directly from the manufacturer to the end customer, bypassing the retailer (who takes the order and initiates the delivery request).
- The retailer carries no inventory.
- Information flows from the customer, via the retailer, to the manufacturer, and product is shipped directly from the manufacturer to customers.

- The biggest advantage of drop-shipping is the ability to centralize inventories at the manufacturer, which can aggregate demand across all retailers that it supplies. As a result, the supply chain is able to provide a high level of product availability with lower levels of inventory.
- Benefit of aggregation is achieved only if the manufacturer can allocate at least a portion of the available inventory across retailers on an as-needed basis. The benefits from centralization are highest for high-value, low-demand items with unpredictable demand.
- Drop-shipping also offers the manufacturer the opportunity to postpone customization until after a customer has placed an order. Postponement, if implemented, further lowers inventories by aggregating to the component level.
- Although inventory costs are typically low with drop-shipping, transportation costs are high because manufacturers are farther from the end consumer
- Supply chains save on the fixed cost of facilities when using drop-shipping because all inventories are centralized at the manufacturer. This eliminates the need for other warehousing space in the supply chain.
- There can be some savings of handling costs as well, because the transfer from manufacturer to retailer no longer occurs.
- A good information infrastructure is needed between the retailers and the manufacturer
- Response times tend to be long when drop-shipping is used because the order must be transmitted from the retailer to the manufacturer and shipping distances are generally longer from the manufacturer’s centralized site.

- Manufacturer storage allows a high level of product variety to be available to the customer.
- Drop-shipping provides a good customer experience in the form of delivery to the customer location.
- The handling of returns is more expensive under drop-shipping because each order may involve shipments from more than one manufacturer.

Cost factor	Performance
Inventory	Lower costs because of aggregation. Benefits of aggregation are highest for low-demand, high-value items. Benefits are large if product customization can be postponed at the manufacturer.
Transportation	Higher transportation costs because of increased distance and disaggregate shipping.
Facilities and handling	Lower facility costs because of aggregation. Some saving on handling costs if manufacturer can manage small shipments or ship from production line.
Information	Significant investment in information infrastructure to integrate manufacturer and retailer.

Cost factor	Performance
Response time	Long response time of one to two weeks because of increased distance and two stages for order processing. Response time may vary by product, thus complicating receiving.
Product variety	Easy to provide a high level of variety.
Product availability	Easy to provide a high level of product availability because of aggregation at manufacturer.
Customer experience	Good in terms of home delivery but can suffer if order from several manufacturers is sent as partial shipments.
Time to market	Fast, with the product available as soon as the first unit is produced.
Order visibility	More difficult but also more important from a customer service perspective.
Returnability	Expensive and difficult to implement.

Manufacturer storage with direct shipping is best suited for:

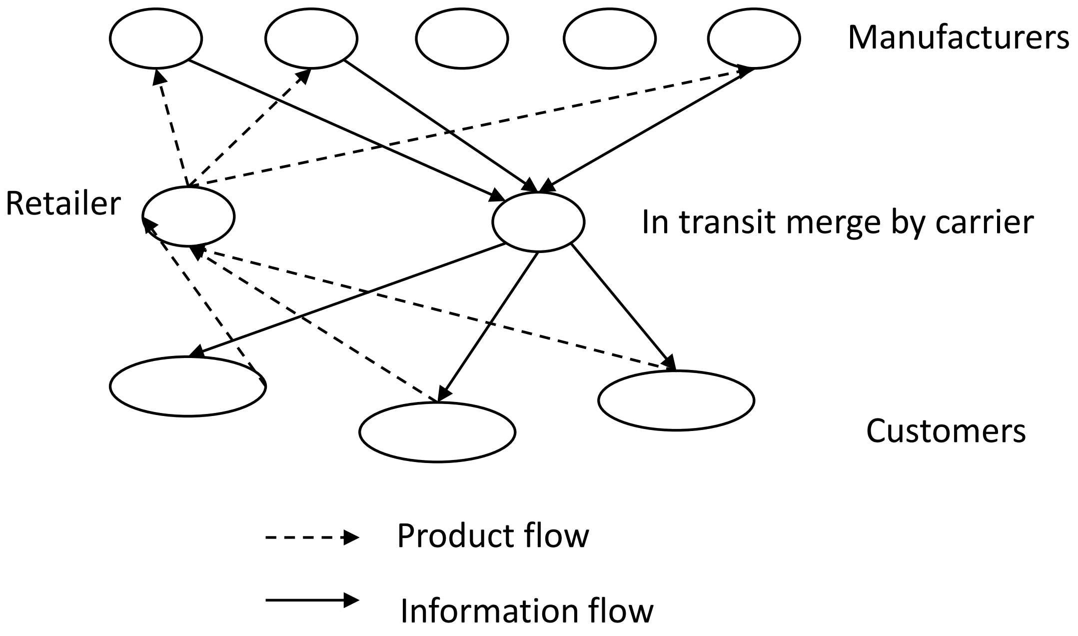
- For a large variety of low-demand
- For high-value items for which customers are willing to wait for delivery and accept several partial shipments. If it allows the manufacturer to postpone customization, thus reducing inventories
- For direct sellers that are able to build to order.

For drop-shipping to be effective, there should be few sourcing locations per order

Manufacturer Storage with Direct Shipping and In-Transit Merge

In-transit merge combines pieces of the order coming from different locations so the customer gets a single delivery.

- This approach has the greatest benefits for products with high value whose demand is difficult to forecast, particularly if product customization can be postponed.
- In-transit merge decreases transportation costs relative to drop-shipping by aggregating the final delivery
- The party performing the in-transit merge has higher facility costs because of the merge capability required. Receiving costs at the customer are lower because a single delivery is received.
- A sophisticated information infrastructure is needed to allow in-transit merge. In addition to information, operations at the retailer, manufacturers, and the carrier must be coordinated.



- Customer experience is likely to be better than with drop-shipping, because the customer receives only one delivery for an order instead of many partial shipments.
- Order visibility is an important requirement.
- Returnability is similar to that with drop-shipping.
- The main advantages of in-transit merge over drop-shipping are lower transportation cost and improved customer experience. The major disadvantage is the additional effort during the merge itself.

Cost factor	Performance
Inventory	Similar to drop-shipping.
Transportation	Somewhat lower transportation costs than drop-shipping.
Facilities and handling	Handling costs higher than drop-shipping at carrier; receiving costs lower at customer.
Information	Investment is somewhat higher than for drop-shipping.

Cost factor	Performance
Response time	Similar to drop-shipping; may be marginally higher.
Product variety	Similar to drop-shipping.
Product availability	Similar to drop-shipping.
Customer experience	Better than drop-shipping because only a single delivery is received.
Time to market	Similar to drop-shipping.
Order visibility	Similar to drop-shipping.
Returnability	Similar to drop-shipping.

Manufacturer storage with in-transit merge is best suited for:

- low- to medium-demand
- high-value items the retailer is sourcing from a limited number of manufacturers.

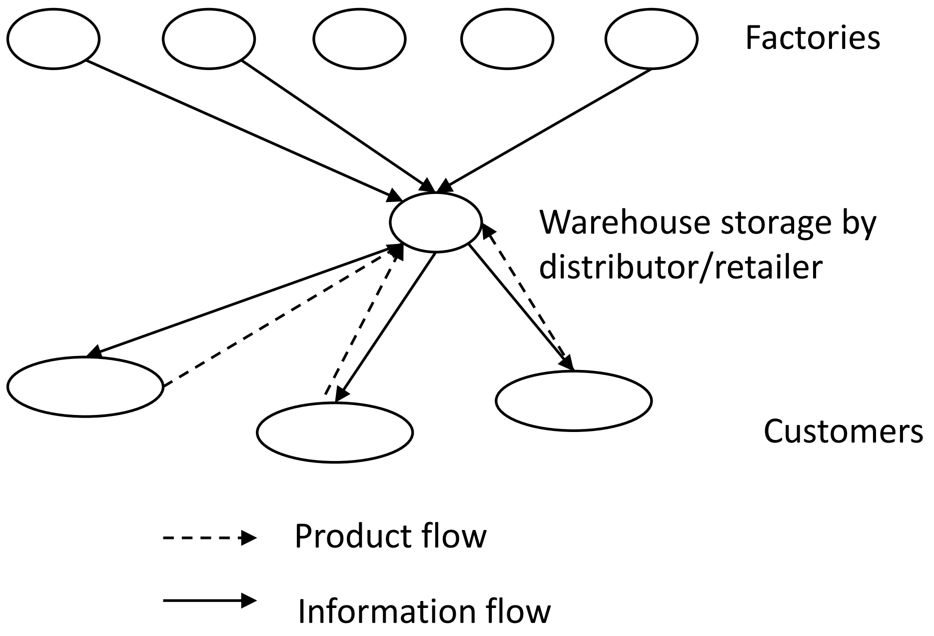
When there are too many sources, in-transit merge can be difficult to coordinate and implement.

In-transit merge is best implemented if there are no more than four or five sourcing locations.

Distributor Storage with Carrier Delivery

Inventory is held not by manufacturers at the factories, but by distributors/ retailers in intermediate warehouses, and package carriers are used to transport products from the intermediate location to the final customer.

- Distributor storage requires a high level of inventory because of a loss of aggregation.
- From an inventory perspective, distributor storage makes sense for products with somewhat higher demand.
- An economic mode of transportation (e.g., truckloads) can be employed for inbound shipments to the warehouse, which is closer to the customer
- Distributor storage allows outbound orders to the customer to be bundled into a single shipment, further reducing transportation cost.
- The information infrastructure needed with distributor storage is significantly less com- plex than that needed with manufacturer storage.



- Real-time visibility between customers and the warehouse is needed, whereas real-time visibility between the customer and the manufacturer is not.
- Response time under distributor storage is better than under manufacturer storage because distributor warehouses are, on average, closer to customers, and the entire order is aggregated at the warehouse before being shipped.

Cost factor	Performance
Inventory	Higher than manufacturer storage. Difference is not large for faster-moving items but can be large for very-slow-moving items.
Transportation	Lower than manufacturer storage. Reduction is highest for faster-moving items.
Facilities and handling	Somewhat higher than manufacturer storage. The difference can be large for very-slow-moving items.
Information	Simpler infrastructure compared with manufacturer storage.

Cost factor	Performance
Response time	Faster than manufacturer storage.
Product variety	Lower than manufacturer storage.
Product availability	Higher cost to provide the same level of availability as manufacturer storage.
Customer experience	Better than manufacturer storage with drop-shipping.
Time to market	Higher than manufacturer storage.
Order visibility	Easier than manufacturer storage.
Returnability	Easier than manufacturer storage.

Distributor storage with carrier delivery is well suited:

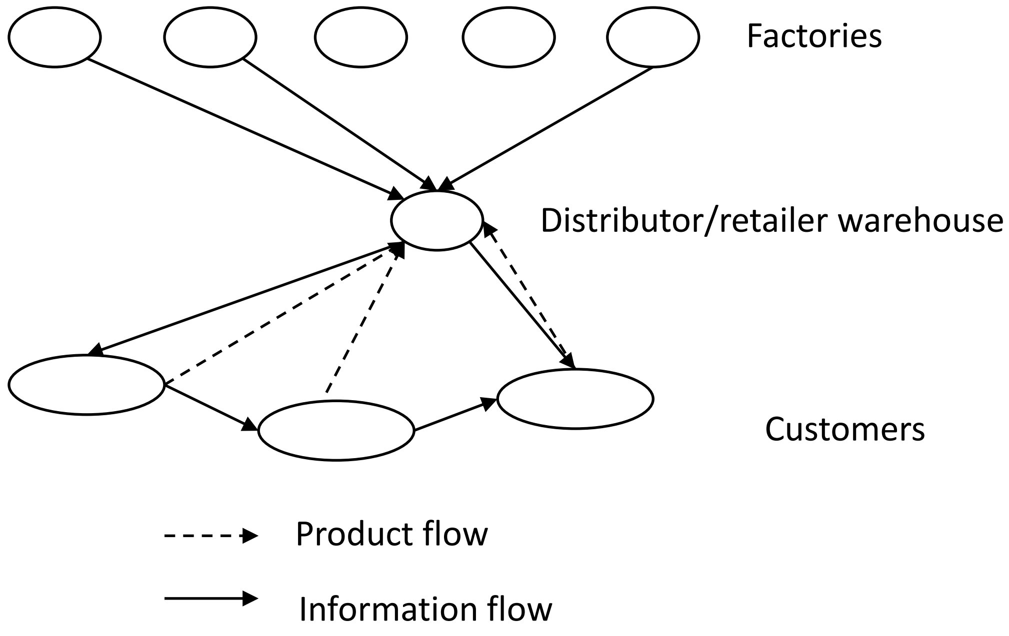
- For slow- to fast-moving items.
- When customers want delivery faster than is offered by manufacturer storage but do not need delivery immediately.

Distributor storage can handle somewhat lower variety than manufacturer storage but can handle a much higher level of variety than a chain of retail stores.

Distributor Storage with Last-Mile Delivery

Last-mile delivery refers to the distributor/retailer delivering the product to the customer's home instead of using a package carrier.

- Unlike package carrier delivery, last-mile delivery requires the distributor warehouse to be much closer to the customer.
- Given the limited radius that can be served with last-mile delivery, more warehouses are required compared to when package delivery is used.
- It requires higher levels of inventory than the other options because it has a lower level of aggregation.
- From an inventory perspective, warehouse storage with last-mile delivery is suitable for relatively fast-moving items that are needed quickly and for which some level of aggregation is beneficial.



- Among all the distribution networks, transportation costs are highest for last-mile delivery, especially when delivering to individuals.
- Last-mile delivery may be somewhat less expensive in large, dense cities, especially if the distributor has very large sales and carries a wide range of products.
- Transportation costs may also be justifiable for bulky products for which the customer is willing to pay for home delivery and where the customer is purchasing in large quantities
- Facility costs are somewhat lower than those for a network with retail stores but much higher than for either manufacturer storage or distributor storage with package carrier delivery.
- The information infrastructure with last-mile delivery is similar to that for distributor storage with package carrier delivery. However, it requires the additional capability of scheduling deliveries. Response times for last-mile delivery are faster than those for package carriers.
- The customer experience can be good
- Returnability is best with last-mile delivery, because trucks making deliveries can also pick up returns from customers.

Cost factor	Performance
Inventory	Higher than distributor storage with package carrier delivery.
Transportation	Very high cost, given minimal scale economies. Higher than any other distribution option.
Facilities and handling	Facility costs higher than manufacturer storage or distributor storage with package carrier delivery, but lower than a chain of retail stores.
Information	Similar to distributor storage with package carrier delivery.

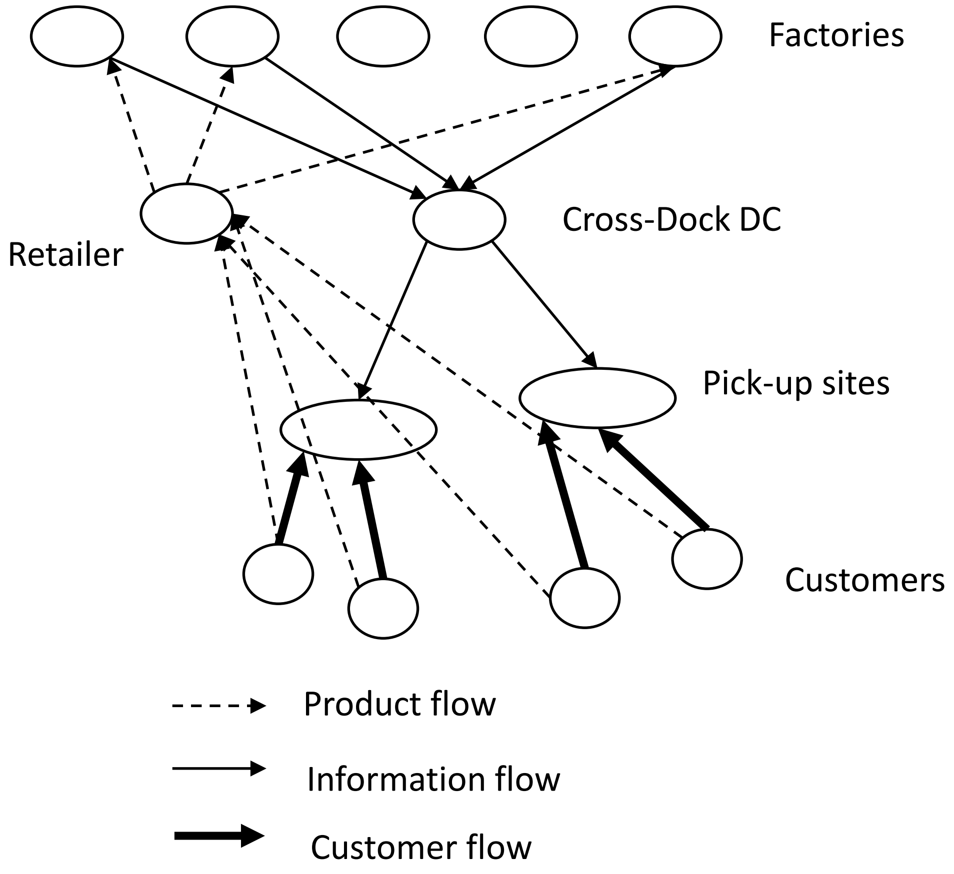
Cost factor	Performance
Response time	Very quick. Same-day to next-day delivery.
Product variety	Somewhat less than distributor storage with package carrier delivery but larger than retail stores.
Product availability	More expensive to provide availability than any other option except retail stores.
Customer experience	Very good, particularly for bulky items.
Time to market	Slightly higher than distributor storage with package carrier delivery.
Order visibility	Less of an issue and easier to implement than manufacturer storage or distributor storage with package carrier delivery.
Returnability	Easier to implement than other previous options. Harder and more expensive than a retail network.

Last-mile delivery may be justifiable if customer orders are large enough to provide some economies of scale and customers are willing to pay for this convenience.

Manufacturer or Distributor Storage with Customer Pickup

Inventory is stored at the manufacturer or distributor warehouse, but customers place their orders online or on the phone and then travel to designated pickup points to collect their merchandise. Orders are shipped from the storage site to the pickup points as needed.

- Inventory costs using this approach can be kept low, with either manufacturer or distributor storage to exploit aggregation.
- Transportation cost is lower than for any solution using package carriers because significant aggregation is possible when delivering orders to a pickup site.
- Facility costs are high if new pickup sites have to be built. A solution using existing sites can lower the additional facility costs.
- Processing costs at the pickup site are high because each order must be matched with a specific customer when he or she arrives.
- Order visibility is extremely important for customer pickups.



Cost factor	Performance
Inventory	Can match any other option, depending on the location of inventory.
Transportation	Lower than the use of package carriers, especially if using an existing delivery network.
Facilities and handling	Facility costs can be high if new facilities have to be built. Costs are lower if existing facilities are used. The increase in handling cost at the pickup site can be significant.
Significant investment in infrastructure is required.	Significant investment in infrastructure is required.

Cost factor	Performance
Response time	Similar to package carrier delivery with manufacturer or distributor storage. Same-day delivery is possible for items stored locally at pickup site.
Product variety	Similar to other manufacturer or distributor storage options.
Product availability	Similar to other manufacturer or distributor storage options.
Customer experience	Lower than other options because of the lack of home delivery. Experience is sensitive to capability of pickup location.
Time to market	Similar to manufacturer storage options.
Order visibility	Difficult but essential.
Returnability	Somewhat easier, given that pickup location can handle returns.

- The main advantages of a network with consumer pickup sites are that it can lower the delivery cost and expand the set of products sold and customers served online.
- The major hurdle is the increased handling cost and complexity at the pickup site.
- Such a network is likely to be most effective if existing retail locations are used as pickup sites, because this type of network improves the economies from existing infrastructure.

Retail Storage with Customer Pickup

Inventory is stored locally at retail stores. Customers walk into the retail store or place an order online or by phone and pick it up at the retail store.

- Local storage increases inventory costs because of the lack of aggregation.
- Transportation cost is much lower than with other solutions because inexpensive modes of transport can be used to replenish product at the retail store.
- Facility costs are high because many local facilities are required.
- Good response times can be achieved with this system because of local storage.
- Product variety stored locally is lower than that under other options.
- Time to market is the highest with this option

Cost factor	Performance
Inventory	Higher than all other options.
Transportation	Lower than all other options.
Facilities and handling	Higher than other options. The increase in handling cost at the pickup site can be significant for online and phone orders.
Significant investment in infrastructure is required.	Some investment in infrastructure required for online and phone orders.

Cost factor	Performance
Response time	Same-day (immediate) pickup possible for items stored locally at pickup site.
Product variety	Lower than all other options.
Product availability	More expensive to provide than all other options.
Customer experience	Related to whether shopping is viewed as a positive or negative experience by customer.
Time to market	Highest among distribution options.
Order visibility	Trivial for in-store orders. Difficult, but essential, for online and phone orders.
Returnability	Easier than other options because retail store can provide a substitute.

- The main advantage of a network with retail storage is that it can lower delivery costs and provide a faster response than other networks.
- The major disadvantage is the increased inventory and facility costs.
- Such a network is best suited for fast-moving items or items for which customers value rapid response.

Selecting a Distribution Network Design

- product characteristics
- network requirements

	Retail storage with customer pickup	Manufacturer storage with direct shipping	Manufacturer storage with direct shipping and in-transit merge	Distributor storage with carrier delivery	Distributor storage with last-mile delivery	Manufacturer/distributor storage with customer pickup
Response time	1	4	4	3	2	4
Product variety	4	1	1	2	3	1
Product availability	4	1	1	2	3	1
Customer experience	1-5	4	3	2	1	5
Time to market	4	1	1	2	3	1
Order visibility	1	5	4	3	2	6
Returnability	1	5	5	4	3	2
Inventory	4	1	1	2	3	1
Transportation	1	4	3	2	5	1
Facility and handling	6	1	2	3	4	5
Information	1	4	4	3	2	5

1 corresponds to the strongest performance and 6 the weakest performance.

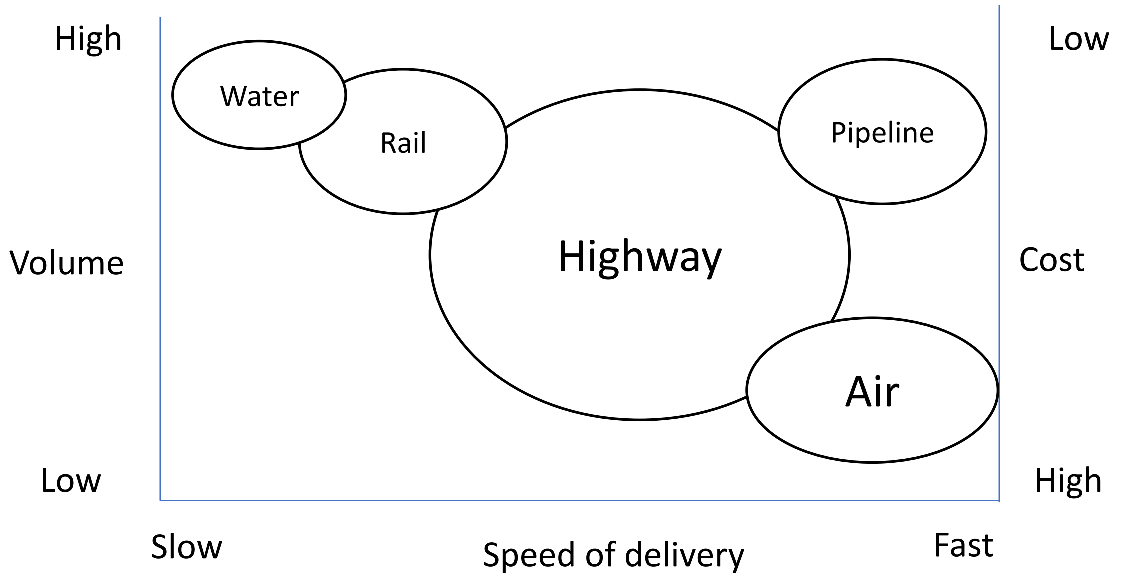
Most companies are best served by a combination of delivery networks. The combination used depends on product characteristics and the strategic position that the firm is targeting

	Retail storage with customer pickup	Manufacturer storage with direct shipping	Manufacturer storage with direct shipping and in-transit merge	Distributor storage with carrier delivery	Distributor storage with last-mile delivery	Manufacturer/distributor storage with customer pickup
High-demandproduct	+2	-2	-1	0	+1	-1
Medium-demand product	+1	-1	0	+1	0	0
Low-demandproduct	-1	+1	0	+1	-1	+1
Very-low-demand product	-2	+2	+1	0	-2	+1
Many product sources	+1	-1	-1	+2	+1	0
High productvalue	-1	+2	+1	+1	0	+2
Quick desired response	+2	-2	-2	-1	+1	-2
High productvariety	-1	+2	0	+1	0	+2
Low customereffort	-2	+1	+2	+2	+2	-1

+2 = very suitable; +1 = somewhat suitable; 0 = neutral; -1 = somewhat unsuitable; -2 = very unsuitable.

Transport modes

- Highway:
  - Truck
  - Parcel
- Rail
- Cargo ship
- Air
- Pipeline



Truck

Size of the product, weight, and liquid or bulk can all be accommodated with this mode.

The amount of cargo that a single truck can carry is limited by size and weight.

Two trucking modes are common:

- Full truckload (FTL):
  - you can choose any origin and destination
  - you generally pay a flat rate per mile
  - the fastest, most reliable way to ship a load
  - more economical for large shipments
- Less than truckload (LTL):
  - your cargo is combined with other shippers' cargo in the same trailer
  - charged only for the space you're using in the trailer (weight and dimensions) and travelled distance
  - better for small shipments

Trucking prices are called rates. Rates are a function of supply, demand, origin and destination, and they change all the time

Parcel

For a relatively small load, such as a couple of boxes.

The price you pay for a parcel shipment depend on weight, dimension, O-D pair, and speed.

Rail

For big, heavy cargo that needs to travel long distances over land.

The suitability of rail can vary depending on the rail infrastructure.

Rail cargo is slow, and transit times can be unpredictable.

Many factories and distribution centers that ship a lot of cargo on trains have installed their own set of tracks, called rail spurs.

Cargo ship

Very high capacity and very low cost, but transit times are slow, and large areas of the world are not directly accessible to water carriers.

Different kinds of ships are designed to carry specific kinds of cargo.

Air

Fast but expensive. Small, light, expensive items are most appropriate for this mode of transportation.

Air cargo is often loaded into special containers called igloos that are designed to protect both the cargo and the airplane.

Pipeline

This is highly specialized and limited to liquids, gases, and solids in slurry forms. No packaging is needed and the costs per mile are low. The initial cost to build a pipeline is very high.

Intermodal

Intermodal transportation is the use of more than one mode of transport to move a shipment to its destination.

Intermodal traffic has grown considerably with the increased use of containers for shipping and the rise of global trade.

For global trade, intermodal is often the only option because factories and markets may not be situated next to ports.

On land, the rail/ truck intermodal system offers the benefits of lower cost than TL and delivery times that are better than rail.

It also creates convenience for shippers that now deal with only one entity representing all carriers that together provide the intermodal service.

Key issues in the intermodal industry involve the exchange of information to facilitate shipment transfers.

Selecting transport modes

Factors:
- Product characteristics: what modes make sense for the product, and the type of container needed
- Facility characteristics: the design and capabilities of the facilities that the product will flow through (origin and destination facilities, distribution centers, ports, and cross-docks)
- Time constraints
- Cost of transport mode: which lane provides the most value for the lowest price

Routing guide: document that specifies rules about which modes to use and which carriers to select based on factors such as the type of cargo, the size of the load, and the O-D pair.

Backhaul: the return movement of a transport vehicle from its original destination to its original point of departure.

Practical challenges:
- Availabilty of the chosen transport mode for an O-D pair
- Availability of the capacity for the chosen transport mode

Introduction

The number and locations of these facilities is a critical factor in the success of any supply chain (80% of the costs of the supply chain).

Aims: to determine the optimal location of facilities (warehouses, plants, lines within the plants, and suppliers) and the best flow of products through this facility network structure:

- How many warehouses should we have, where should they be, how large should they be, what products will they distribute and how will we serve our different types of customers?
- How many plants or manufacturing sites should we have, where should they be, how large should they be, how many production lines should we have and what products should they make, and which warehouses should they service?
- Which products should we make internally and which should we source from outside firms?
- If we source from outside firms, which suppliers should we use?
- What is the trade-off between the number of facilities and the service level? How much does it cost to improve the service level?

Factors affecting the location decision

Transportation Cost
The location of your facilities determines the distance you need to move product, which directly impacts the amount you spend on transportation.

Service Level
Proximity to customers impacts the time it takes to get product to your customers.

Risk
The number and location of facilities impacts risk for the production in case of disruptions. Political risk should be considered.

Local Labor, Skills, Materials, and Utilities
The location of facilities also determines what is paid for labor, findingneeded skills, the cost of locally procured materials, and the cost of utilities.

Taxes
Depending on where facilities are located, the type of operations being performed and shipping products.

Infrastructures

Adequate transport infrastructures, energy and telecommunications requirements and government’s willingness to invest in upgrading infrastructure.

Suppliers

A high-quality and competitive supplier base makes a given location suitable

Other Facilities

The location of other plants or distribution centers of the same company impact product mix and capacity.

Carbon Emissions

Locating facilities to minimize the distance traveled or the transportation costs often has the side benefit of reducing carbon emissions

Environmental Regulation

Besides measurable cost implications, these regulations influence the relationship with the local community

Roles and types of warehouses and plants

A warehouse represents a facility where firms store product or a location where product simply passes through from one vehicle to another. Warehouses are needed in a supply chain for the following reasons:

- Consolidation of Products: bringing different products together to make a single shipment to a customer. This will be cheaper than having the products ship to the customers directly from each individual source of supply.
- Buffer Lead Time: shipping to your customers with lead time that is shorter than that which can be offered by shipping directly from the plant or supplier location
- Service Levels: the need to be close to customers can create the need for multiple warehouses
- Production Lot Sizes: production plans attempt to maximize the number of units of product made during each run (lot size). This requires the extra units to be "stored" in warehouses until future demand requires them
- Inventory Pre-Build: some industries see huge spikes in the supply of raw materials or in the demand of finished goods. These abundant raw materials must be stored until the time they will be needed for steady monthly production cycles. Additional capacity during off-peak time periods is used to store finished-good units awaiting their use to fulfill the upcoming spikes in demand
- Transportation Mode Trade-offs: a warehouse can help reduce costs by allowing the shipment of products a long distance with an efficient (and lower cost) mode of transportation and then facilitating the changeover to a less efficient (and usually more expensive) mode of transportation for a shorter trip to the final destination (as opposed to shipping the entire distance on the less efficient mode).

A supply chain may have many types of warehouses to meet many different needs:

- Distribution Center: a warehouse where product is stored and from which customer orders are fulfilled
- Cross-Dock: a warehouse that is simply a meeting place for products to move from inbound trucks to outbound trucks
- Plant-Attached Warehouse: a warehouse that is attached to a manufacturing plant (at the end of the line where product is staged prior to being loaded onto a truck for shipment or a storage point for product made at the plant or for products made at other plants). A major benefit of a plant- attached warehouse is the reduction of transportation costs because a product does not have to be shipped to another location immediately after it comes off the end of the line
- Hub Warehouse or Central Warehouse: a warehouse that consolidates products to be shipped to other warehouses in the system before moving on to customers. Different from cross-docks, the products are normally stored in these locations for longer periods of time before being used to fulfill demand

There are also needs for different temperature classes (frozen, refrigerated, or ambient), different levels of safety (hazardous or nonhazardous), and different levels of ownership (company owned, company leased, or the company uses a third-party facility.

A plant is a location where product is made or where it comes from. It could be a manufacturing plant that produces raw material, components, or finished goods, or just does assembly.

The location of the plant can impact transportation costs and the ability to service customers. Factors that lead to multiple plants making the same product:

- Service Levels. Proximity of plants to customer affects service level (especially if warehouses are not used)
- Transportation Costs, in particular for producers of heavy or bulky products that easily fill up truckload capacities
- Economies of Scale. The higher the production level at a single location, the lower the production cost per unit
- Taxes
- Steps in the Production Process. For example, it can often be a good strategy to make product in bulk at a low-cost plant, ship it in bulk to another plant closer to the market to complete the conversion to a finished good

A plant may represent:

- A Manufacturing or Assembly Site: a site that is owned by the firm that makes products
- A Supplier: a location that is not owned by the firm but supplies product to the firm
- Third-Party Manufacturing Site: a location similar to supplier plants, but these sites make product on behalf of the firm and are therefore treated more like the firms' own plants than a raw material supplier

Elements of an optimization problem

Mathematical optimization technology is the best way to sort through the various options, balance the trade-offs, determine the best locations for facilities, and support better decision making.

Optimization is a complementary, not competitive, technology that allows you to actually determine the best locations for your facilities.

To formulate a logical supply chain network model, the following four elements are needed:

- Objective. The objective is the goal of the optimization and the criteria we'll use to compare different solutions (e.g. minimizing cost). An optimization problem needs to have a quantifiable objective.
- Constraints. The constraints define the rules of a legitimate solution (e.g. which products may be made where, how much production capacity is available, how close your warehouses must be to customers)
- Decisions. The decisions (decision variables) define what you allow the optimization to choose from (e.g. where those sites are, and what product is made in which location). The allowable decisions cannot be separated from the constraints.
- Data

The problem is translated into a series of equations and then solved using linear and integer programming techniques.

Mixed Integer Programming: a series of steps that are influenced by the objective, the constraints, and the decision variables. During the steps, the objective steers the solution to more favorable costs and avoids less favorable costs, the constraints set the rules and can prevent it from doing more of what it wants or can force it to do something that is not favorable to the objective, and the decision variables tell it what it is allowed to change.

The role of qualitative data

Nonquantifiable factors:

- Firm’s Strategy. The firm may value cost more than service or vice versa
- Risk
- Disruption Cost. changes could cause significant disruption
- Willingness to Change
- Public Relations and Branding. Firms with a highly visible brand need to consider the public reaction and the impact on their brand
- Competitors
- Union versus Non-Union
- Tax Rebates
- Relationships with Trucking Companies, Warehousing Companies, and Other Supply Chain Partners

Main challenges in location planning

A supply chain study must span many different areas of an organization: sales, operations, logistics, finance, and IT.

1. To understand and start to balance the different and conflicting objectives that each group may have:
    - Sales Team— create many warehouses and have small frequent shipments to customers.
    - Operations Team (Production)— produce product in one location to maximize economies of scale.
    - Operations Team (Warehousing)—minimize warehousing locations to reduce fixed and management costs.
    - Logistics Team (Transportation)—have large shipments on less costly modes of transportation.
    - Finance Team— have low levels of inventory and operations requiring the least investment in warehousing and production locations, and incur the lowest costs tied to logistics.
2. To collect and validate data from all these different parts of the organization, and to estimate data for the new potential locations and product flow paths
3. To actually implement the results, since changes to a supply chain may also cause a temporary state of disruption

Basic model: the Center of Gravity problem

The simplest facility-location problem is the center of gravity (COG) problem. For logistics, a center of gravity problem is usually defined as selecting the location of a facility so that the weighted-average distance to all the demand points is minimized.

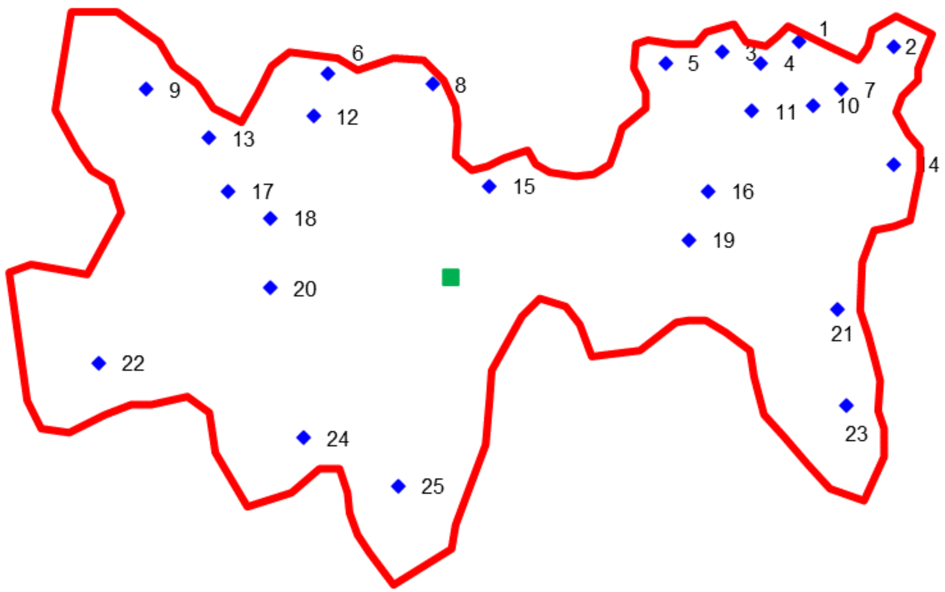
A center of gravity solution suggests that facilities are located at the center (the “center of gravity”) of a collection of demand points (or in some instances, for firms with many suppliers, at the center of the supply points).

Center of gravity models are, by definition, clear-cut and not ambiguous.

Problem 1: Physics Weighted-Average Centering

The citizens of Logistica must choose a location for their capital.

They choose their capital location by considering the boundaries, and then selecting the location that centers the country geographically.



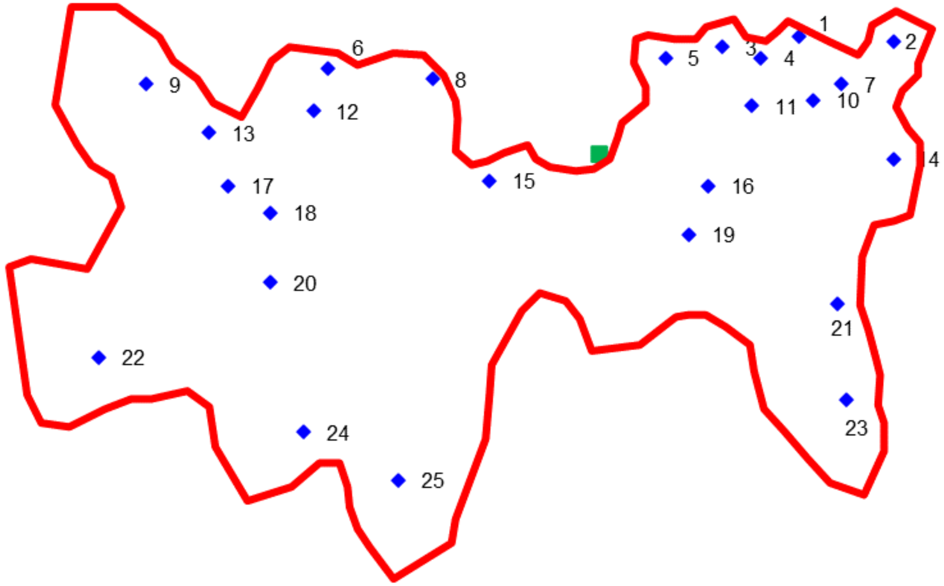
City	Population	Distance to Capital (Miles)
1	1'655'000	588
2	2'300'000	707
3	601'000	489
4	1'385'000	527
5	1'230'000	415
6	665'000	318
7	664'000	615
8	885'000	250
9	1'116'000	503
10	636'000	569
11	1'200'000	484
12	148'000	288
13	854'000	395
14	1'473'000	658
15	615'000	130
16	1'145'000	389
17	627'000	343
18	542'000	273
19	379'000	348
20	964'000	263
21	546'000	560
22	706'000	522
23	727'000	596
24	669'000	298
25	931'000	280
Weighted Average Distance:		471

Then they decide to locate the capital based on a weighted average of locations of the cities already scattered about the country. This physics center of gravity problem can be solved with a simple closed-form equation finding the coordinates of Logistica's capital:

$$LON_{cap} = \frac{\sum_c p_c LON_c}{\sum_c p_c} \qquad LAT_{cap} = \frac{\sum_c p_c LAT_c}{\sum_c p_c}$$

LON represents a city's longitude, LAT represents its latitude, p represents a city's population, c is the set of cities.
The population is used as the weighting factor. In network problems, customer demand is the most common weighting factor.

Facility location planning



Though it's good for a physics calculation, the planners quickly realize that the center of gravity calculation would not prove practical.

City	Population	Distance to Capital (Miles)
1	1'655'000	327
2	2'300'000	452
3	601'000	226
4	1'385'000	265
5	1'230'000	157
6	665'000	409
7	664'000	363
8	885'000	260
9	1'116'000	662
10	636'000	318
11	1'200'000	229
12	148'000	418
13	854'000	566
14	1'473'000	428
15	615'000	162
16	1'145'000	164
17	627'000	540
18	542'000	482
19	379'000	167
20	964'000	504
21	546'000	395
22	706'000	770
23	727'000	479
24	669'000	558
25	931'000	511
Weighted Average Distance:		388

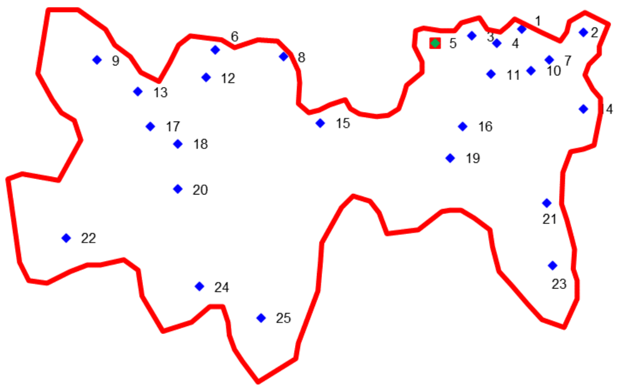
When finding the center of gravity, the principles of physics would naturally square distance. The formula does not minimize the weighted average distance to each city. Instead, the formula minimizes the demand multiplied by the distance squared. However, to real world supply chain practitioners, this definition does not necessarily make sense.

In summary, the physics center of gravity calculations should not be used for the following reasons:

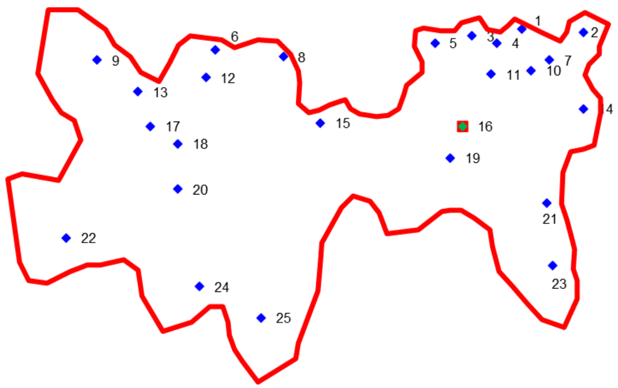
- It does not minimize the weighted average distance.
- It might point to anywhere in space.
- It will almost never land in a location that exploits existing infrastructure.
- The method cannot take advantage of true road distances (it has to rely on straight-line estimates based on latitude and longitude).
- It cannot be extended to include factors like costs, capacities, different types of facilities, different products, multiple levels of facilities (hub and spoke, suppliers, warehouses, retailers, etc.), or other practical considerations.

Problem 2: Practical Center of Gravity

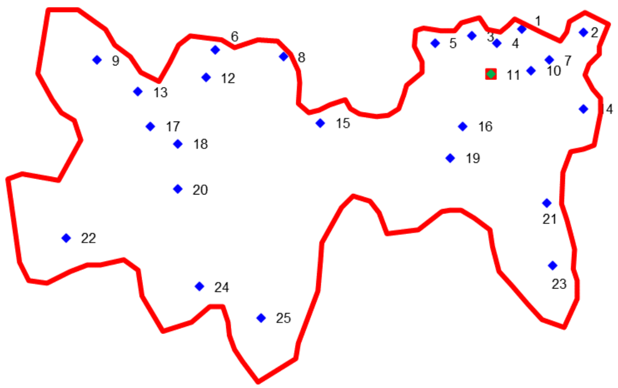
Try to find an approach to minimize the average distance traveled



City 5



City 16



City 11

	Center of Gravity	City 5	City 16	City 11
Weighted Average Distance	388	380	378	370
% within 100 miles	0%	8%	7%	17%
% within 200 miles	15%	32%	29%	39%
% within 300 miles	33%	39%	48%	58%

- The physics center of gravity solution did not do very well on the key statistics for judging potential locations. It had a higher average distance and had very few citizens with 100, 200, or 300 miles radius (because the objective of the physics center of gravity does not minimize average distance or percentage of customers close to the capital)
- There may be trade-offs to make (e.g. it may be important that a large portion of the population can easily get to the capital city, even if the average distance is a bit longer)

Because the planners are picking only one city as the capital, they can easily calculate the key statistics for all 25 cities.

	Center of Gravity	City 4	City 10	City 15	City 20	City 22
Weighted Average Distance	388	381	399	428	585	800
% within 100 miles	0%	24%	24%	3%	7%	3%
% within 200 miles	15%	48%	49%	7%	12%	3%
% within 300 miles	33%	56%	58%	12%	23%	10%

A more realistic problem selects the most central location from a list of possible candidates, if there is time to enumerate all the solutions like the planners of Logistica did. This more realistic and practical formulation of the problem is best solved with linear and integer programming techniques.

Distance-based approach

Aim: optimally selecting “P” facility location(s) considering only distance.

Retail Case Study: Al’s Athletics
(Watson et al., 2013)

Aim: selecting optimal “P” warehouses from a predetermined list of options

Formal Problem Definition for Locating “P” Facilities: “Given a set of customer locations and their demand, find the best P number of facilities (plants or warehouses) that minimize the total weighted distance from the facility to the customer, assuming that each facility can satisfy the full demand of the customer and that all demand is always satisfied.”

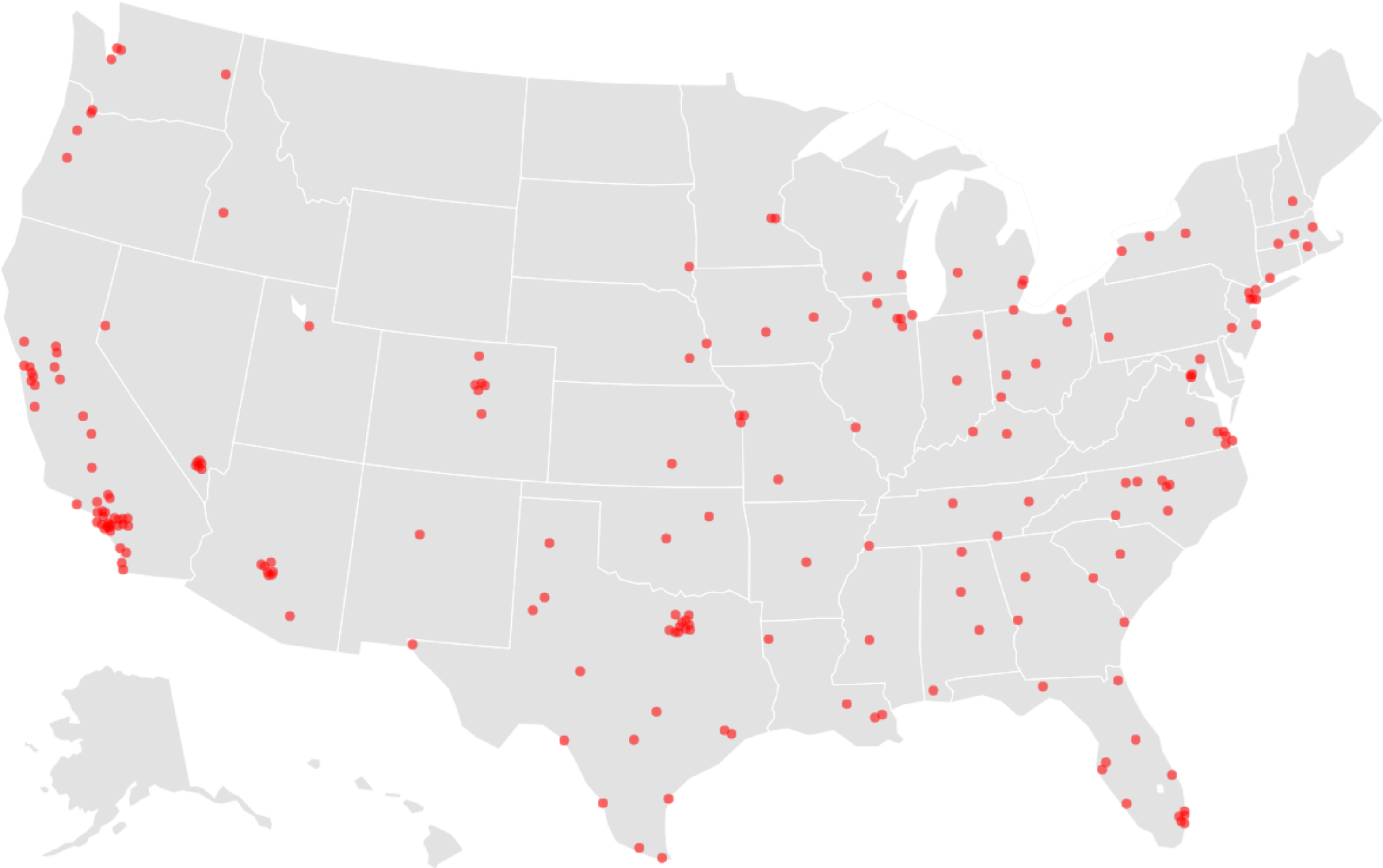
Customers

When modeling a supply chain, the customer refers to the final delivery location for our products. We are only concerned with the final delivery point that is relevant for the supply chain we are analyzing. A customer location may represent:

- In a retail supply chain, each of the store’s locations (as in the Al’s case study)
- If the retailer sells directly to their online customers and delivers to the home, each ZIP or postal code they ship to
- If the retailer has stores in a shopping mall, they may often ship to the store in a truck that is shared with other retailers in the same mall. The common destination is called pool point and represents the customer location.

- If there is a manufacturing firm that sells to retailers or wholesalers, the customers could be the warehouses of the retailer or wholesaler.
- If the manufacturing firm ships to other manufacturing firms, the customers will be the other manufacturing sites.
- If a firm’s supply chain already has a fixed and unchanging set of warehouses with a fixed set of customers they service, these warehouses are the customer points.
- If a firm’s supply chain services the home-building industry, the customer may represent the job site where the firm need to provide the product
- If a company’s supply chain exports product and it is not responsible after the product leaves the country, the customer can be the port of exit.

Facility location planning



Demand

The demand is used to determine the relative importance of each customer.
An appropriate unit of measure is to be specified, that correctly apply the associated manufacturing and logistics costs and capacities.
Common units of measure include:

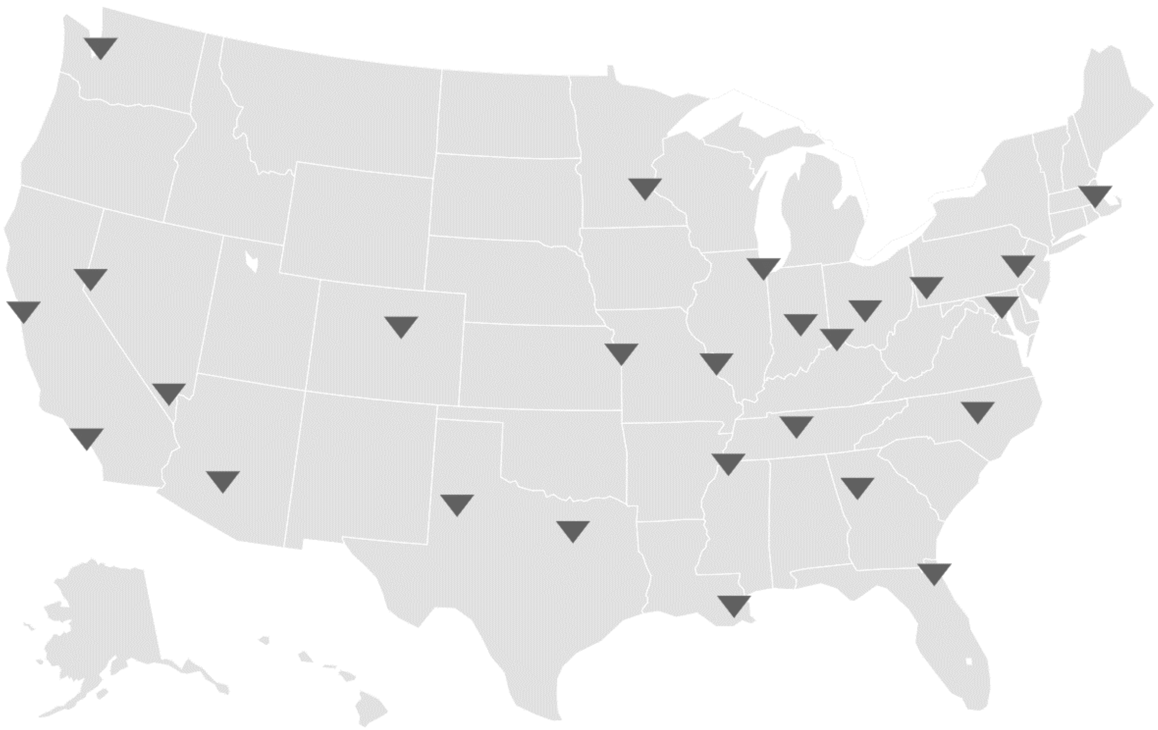
- Total weight
- Total pallets
- Total cube (physical space the product occupies)
- Total truckloads
- Total cases

Al's selection of weight is probably the most common unit of measure in network design due to the fact that most transportation costs are based on the weight.

Time horizon for demand analysis: Typical network design studies use an entire year's worth of demand; reviewing an entire year of activity within a supply chain ensures that all peaks and valleys of your customers' buying patterns are accounted for in your analysis

Problem definition: “Find the best P number of facilities (warehouses) that minimize the total weighted distance from the facility to the customer, assuming that each facility can satisfy the full demand of the customer and that all demand is satisfied.”

- The measure of “best” is that we are minimizing the weighted-distance to customers
- Facilities are picked from a predefined set of locations



- List of the top 25 most frequently selected locations for warehouses
- Matrix with the distance between each facility and each customer (road distance)

In the Logistica example, we saw how finding the single-best point was relatively easy. We simply listed all the possible combinations and picked the best one.
For Al's problem, 26 choices are possible. However, the complication lies in the number of potential combinations we must consider to prove this

Total potential facilities	**25**
Number picked	**Total combinations**
2	300
3	2300
4	12650
5	53130
6	177100
7	480700
8	1081575
9	2042975
10	3268760

Total potential facilities	**250**
Number picked	**Total combinations**
2	31125
3	2573000
4	158882750
5	7817031300
6	319195444750
7	11126241217000
8	337959576966375
9	9087357513984750
10	219005316087033000

The number of combinations explodes as more potential facilities are considered. Linear and integer programming techniques are needed for sorting through these combinations in a systematic fashion to find the best answer.

Indexes:

- Set of customers to serve: j=New York, Chicago, …, J
- Demand of each customer d_j
- Set of potential facilities we can select from: i= warehouse 1, warehouse 2, …, I
- Distance from facility i to customer j: $dist_{i,j}$

Objective: Minimize the average weighted distance from the warehouses to the stores.

Constraints:

- All the demand must be met
- The number of facilities must be limited to P

Decisions:

- Is the warehouse at location i selected? (If so, X_i=1, otherwise, 0)
- Does facility i service customer j? (If so, Y_i=1, otherwise, 0)

The objective function determines whether one solution is better than another. It minimizes the total weighted distance from warehouses to customers:

$$\sum_{i \in I} \sum_{j \in J} dist_{i,j} d_j Y_{i,j}$$

Constraints:

1. Every customer must be fully served:

$$\sum_{i \in I} Y_{i,j} = 1 \ \forall j \in J$$

2. Exactly P facilities must be located:

$$\sum_{i \in I} X_i = P$$

3. If a warehouse serves a customer, then that warehouse must be considered “selected”:

$$Y_{i,j} \leq X_i \ \forall i \in I, \forall j \in J$$

4. X and Y must be 0 or 1:

$$Y_{i,j} \in \{0,1\} \forall i \in I, \forall j \in J$$

$$X_i \in \{0,1\} \forall i \in I$$

$$min \sum_{i \in I} \sum_{j \in J} dist_{i,j} d_j Y_{i,j}$$

$$\sum_{i \in I} Y_{i,j} = 1 \ \forall j \in J$$

$$\sum_{i \in I} X_i = P$$

$$Y_{i,j} \le X_i \ \forall i \in I, \forall j \in J$$

$$Y_{i,j} \in \{0,1\} \forall i \in I, \forall j \in J$$

$$X_i \in \{0,1\} \forall i \in I,$$

The problem can be solved for any given set of data.

There are a lot of different combinations of potential warehouses. What makes the number of combinations difficult is that there are no reliably efficient techniques for finding the optimal solution when you have binary variables that take on either a zero or a one.

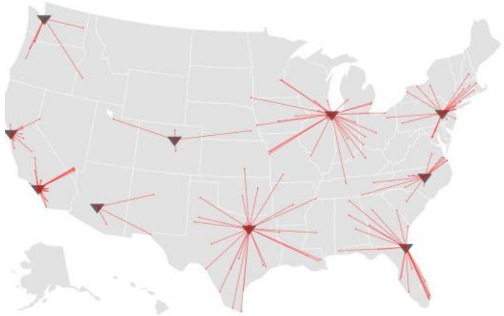
NP-hard problem: there are no known algorithms that are guaranteed to solve the problems to absolute optimality in a reasonable time as the model size grows

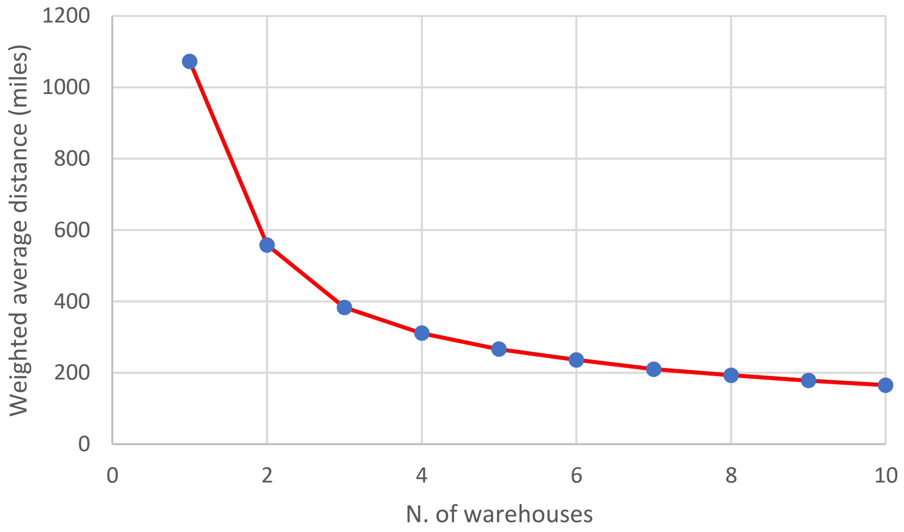
The best value for P is unknown. Multiple scenarios with different values for P are run and results are compared.


P = 1


P = 2


P = 5


P = 10



- There are diminishing returns to adding more facilities
- Weighted average distance is a good summary of service level

Percentage of demand served within certain distance bands or ranges:

Distance band (miles)				
N. of warehouses	100	400	800	3200
1	1	4	30	65
2	12	23	39	26
3	18	38	35	9
4	31	38	24	7
5	36	38	23	3
6	41	36	22	1
7	45	39	17	1
8	47	38	15	0
9	49	38	13	0
10	51	40	9	0

- Solutions improve within each distance band
- These measures do not necessarily have to improve because they are not part of the objective
- These distance bands are closely correlated with improving the average distance
- It is important to run multiple scenarios to understand the trade-offs, and understand the marginal value of adding facilities

Modelling service level

Service level may mean:

- Minimize the average distance
- Maximize the percentage of customers within a certain distance

Other measures of service level not inputs to an accurate network design model:

- Fill Rate: percentage of orders that are filled from inventory
- Late Orders: how late the shipment is (when a customer orders a product, when the firm promises a delivery, and whether that delivery actually happens on time)

Case Study: Chen's cosmetics
(Watson et al., 2013)

Aim: adding two production facilities in locations across China ensuring that these additional locations offer the best service to Chen's existing customer base.

Service levels within network design are most commonly measured by transit time or distance. Gathering data on transit times for all existing and potential lanes is not always an easy task for companies, however. Therefore, network design analysts commonly use a distance equivalent to transit time.

The number of kilometers that could be covered within a day would be limited:

- Delivery trucks travel at an average speed of only 50kmph
- Drivers can only work a maximum of eight hours per day
- One-day transit distance: 400 km
- Desired two- day service level is a distance band of 800km from plants outbound to distributor warehouse locations

Facility location planning

Scenarios:

- As-is network
- Objective 1: minimizing average distance (offer the best service possible to customers who purchase the most)
- Objective 2: maximizing the demand within 800km (2 days service)

	As-is network	Objective 1	Objective 2
Average distance (km)	1603	588	654
% demand within 2 days service	18	78	80

Offering the best service to the highest-demand customers has little to no effect on the ability to maximize service to markets within 800km, and therefore these locations are ideal for the needed expansion.

Objective function that maximizes the amount of demand with a certain distance:

$$\max \sum_{i \in I} \sum_{j \in J} (dist_{i,j} > HighServiceDist ? 0 : 1) d_j Y_{i,j}$$

If the distance between the two points is greater than the service parameter, the expression takes on a value of 0, otherwise, 1.

- This guides the objective to look for as many combinations as possible in which the demand is within the 800km
- No guidance is provided for how to assign customers outside the HighServiceDist, so, it will make a random selection. Therefore, a constraint is needed to tell the solver engine that even if a customer cannot be assigned to a facility within the HighServiceDist restriction, it should be assigned to a facility that is reasonably close:

$$\sum_{i \in I} \sum_{j \in J} dist_{i,j} d_j Y_{i,j} < AvgSertviceDist \sum_{j \in J} d_j$$

The AvgServiceDist constant represents the largest average customer demand assignment distance that will be tolerated.

$$\max \sum_{i\in I}\sum_{j\in J}(dist_{i,j} > HighServiceDist ? 0:1) d_j Y_{i,j}$$

$$\sum_{i\in I}\sum_{j\in J} dist_{i,j} d_j Y_{i,j} < AvgSertviceDist \sum_{j\in J} d_j$$

$$\sum_{i\in I} Y_{i,j} = 1 \ \forall j \in J$$

$$\sum_{i\in I} X_i = P$$

$$Y_{i,j} \leq X_i \ \forall i \in I, \forall j \in J$$

$$Y_{i,j} \in \{0,1\} \forall i \in I, \forall j \in J$$

$$X_i \in \{0,1\} \forall i \in I$$

The MIP engine might still assign a few stray demand points to unreasonably distant servicing facilities, while nevertheless maintaining a reasonable average servicing distance across the supply chain as a whole (for demand points that are fairly small).

To address this problem, it is common to apply constraints that limit the maximum servicing distance (customer j cannot be assigned to facility i if the distance from i to j is larger than the MaximumDist constant value):

$$Y_{i,j} \leq \left(dist_{i,j} > MaximumDist ? 0:1\right) \forall i \in I \ \forall j \in J$$

MaximumDist should be greater than AvgServiceDist:

- HighServiceDist = 800km
- AvgServiceDist = 1000km
- MaximumDist = 1200km

Model 1:
- Objective function that maximizes the amount of demand with a certain distance (HighServiceDist)
- Constraint for largest average customer demand assignment distance that will be tolerated (AvgServiceDist)

↓

- 75% of demand to a facility within 800km
- Average servicing distance no greater than 1,000km

Model 2:
- Objective function that minimizes the average servicing distance (AvgServiceDist)
- Constraint for the quantity of demand met with high service (HighServiceDist)

↓

- Average servicing distance of 950km
- 75% of demand with a servicing facility within 800km

- It might be possible for a solution to meet more than 75% of demand with high service, but only by sacrificing the average servicing distance and allowing it to go higher than 950km (indeed, higher than 1,000km).
- It might also be possible for a solution to achieve an average servicing distance lower than 950km, but only by allowing more than 25% of demand to be met through a low-service distance.

If a solution has this multiple-goal property, under which one cannot improve one objective without degrading the other, it can be called Pareto optimal.

Model 1

$$\max \sum_{i \in I} \sum_{j \in J} (dist_{i,j} > HighServiceDist ? 0 : 1) d_j Y_{i,j}$$

$$\sum_{i \in I} \sum_{j \in J} dist_{i,j} d_j Y_{i,j} < AvgSertviceDist \sum_{j \in J} d_j$$

Model 2

$$\min \sum_{i \in I} \sum_{j \in J} dist_{i,j} d_j Y_{i,j}$$

$$\sum_{i \in I} \sum_{j \in J} (dist_{i,j} > HighServiceDist ? 0 : 1) d_j Y_{i,j} \geq HighServiceDemand$$

This value for this constraint can be determined from the result of the first model.

$$Y_{i,j} \leq (dist_{i,j} > MaximumDist ? 0 : 1) \forall i \in I \ \forall j \in J$$

$$\sum_{i \in I} Y_{i,j} = 1 \ \forall j \in J$$

$$\sum_{i \in I} X_i = P$$

$$Y_{i,j} \leq X_i \ \forall i \in I, \forall j \in J$$

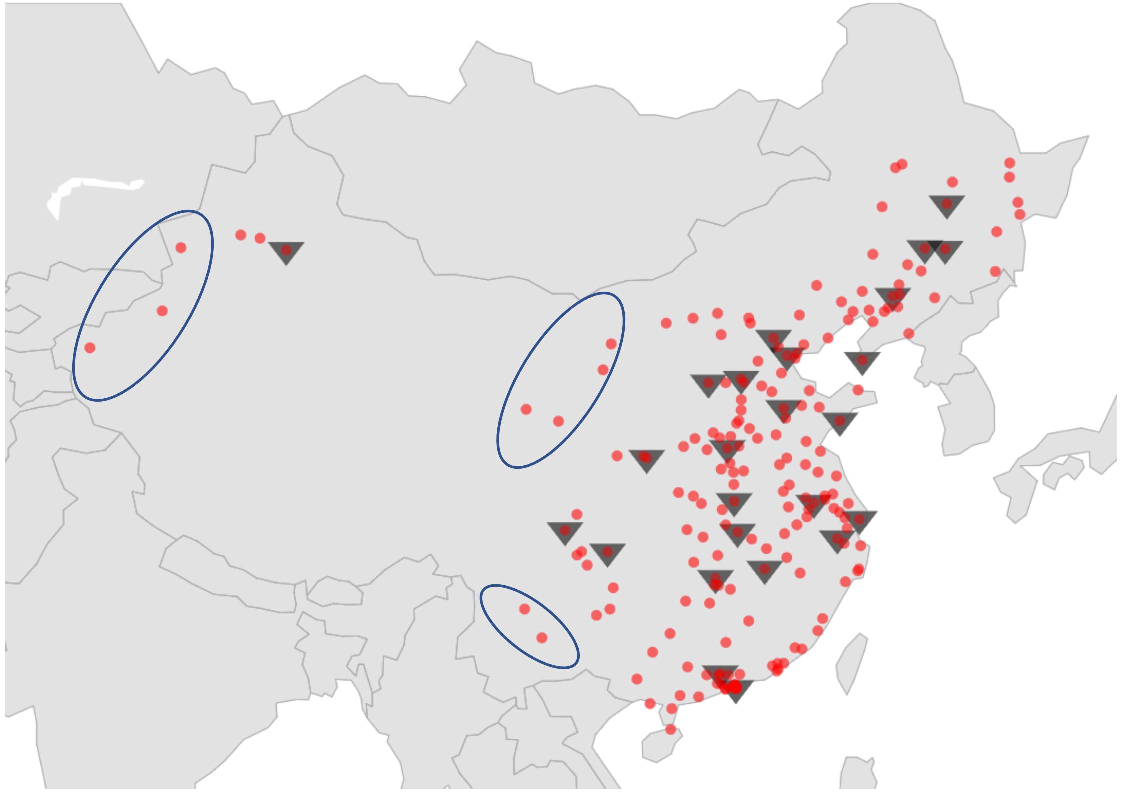
$$Y_{i,j} \in \{0,1\} \forall i \in I, \forall j \in J$$

$$X_i \in \{0,1\} \forall i \in I$$

To consider service level as just a part of their entire network optimization, the service level is no longer the overall solver objective but a constraint within the larger model.

A modeler on Chen's team thinks it will be best for business if the solution ensures that all customers can be serviced within one-day transit or 400km. After quick analysis of their data, it is clear that this constraint will cause infeasibility in their model.

Distributor location	Distance from the closest potential plant
City 1	1076
City 2	694
City 3	669
City 4	62
City 5	536
City 6	521
City 7	514
City 8	511
City 9	505
City 10	499
City 11	476
City 12	464
City 13	464
City 14	456
City 15	444
City 16	433
City 17	427



The best that Chen's can ensure with their current set of potential plant locations is a maximum of three-day transit to all customers. Therefore, the solver must select plants in locations where all distributors are no more than 1,200km away. After this constraint is applied, however, the model run resulted in no feasible solution.

When infeasible solutions are determined, the most common way to pinpoint their cause is to test iterations of providing the model slightly more freedom in its decisions until a feasible solution is possible:

Maximum distance constraint	Feasible solution
1200km (original constraint)	No
2000km	Yes
1800km	No
1600km	Yes

The infeasibility occurs when the model is asked to select plant locations that result in the ability to service all customers within a max of 1,600km or less. Based on the previous discovery that all customers are located within 1,076km of at least one potential plant, it's the combination of the constraints that causes the infeasibility.

Increasing the maximum number of plants that may be selected to four units leads to a feasible solution, confirming that it is the combination of the number of plants and the distance restriction that is causing the problem.

To manage infeasibility:

- Constraints are often altered to incorporate only a portion of the population of the model. In the example, the model may be altered to constrain the resultant service level so that at least 75% of customers are within 1,200km.
- Subgroupings of the customers are created based on their importance. This might result in Tier 1, Tier 2, and so on, designations applied to each. Then the modeler may either apply different levels of constraints to each or apply constraints to only some of the subgroupings. In the example, Chen’s may only want to ensure that their Tier 1 customers are within a maximum of 1,200km but the model has no restrictions on servicing the other locations.

Sensitivity analysis

Network design models may make assumptions in future demand, transport costs, labor costs. It is important that any and all solutions are tested for robustness (sensitivity analysis). By testing changes in key variables such as demand, transportation costs, and labor costs, a modeler is ensuring that an assumption made in one model input is not dramatically altering the resultant savings and network recommendations within the solution.

Modeling Service Levels as Model Objectives:

- Minimize the average distance: Selects optimal network structure in order to minimize the weighted distance traveled or, simply stated, asks the model to service the highest levels of demand within the shortest distances.
- Maximize the percentage of customer or demand within a certain distance: Maximizes customers or demand serviced within a maximum of x distance or, simply stated, asks the model to maximize the number of customers or percent of demand that can be serviced within a distance determined to be an acceptable service level.

Modeling Service Levels as Constraints:

- Maximum Average Distance Constraints: Helps guide the model to return solutions with reasonable assignments for those customers that cannot be reached within a pre-defined service territory.
- Maximum Distance Constraints: Forces the model to service all customers within a specified distance.
- Maximum Distance Constraints per Customer Subsets: Tailors the previous Max Distance Constraint to a specific subset of customers within the model.

Network design models considering capacity

Warehouses have three primary functions: receive, ship product, and store product:

- The capacity for shipping and receiving depends on how many shifts you run, how many people or resources (like automated pickers) you have, what kind of equipment you have, and how much space you have dedicated to these functions.
- The storage capacity is measured by how much physical space you have and how much is in the warehouse at any given time. However, this is not easy to measure.

A warehouse capacity utilization of over 100% is not violating some rule of mathematics. Instead, from a business point of view, it means that the warehouse is very busy.

Manufacturing capacity is typically more straightforward to measure than warehousing capacity. And it is usually more expensive to add. It can be defined by determining how much product can be produced over the year for a given number of shifts

- For warehouse capacity constraint is real, the cost of new warehouse space should be considered:
  - For third-party space, adding capacity may just incur extra variable costs.
  - If your warehouse systems are relatively simple and manual, adding space may be no problem at all. (e.g. by adding racks to ceilings).
  - For a high-speed, highly automated warehouse, more significant capital investment are needed
- If the warehouse constraint is not in space, but in terms of its ability to process inbound and outbound shipments, increasing capacity may require additional shifts, additional dock doors, or more automation.
- For plants, capacity expansion typically comes in three forms:
  - Adding Labor. Extra labor can provide the ability to produce more.
  - Adding Shifts. When you add shifts, additional fixed and variables costs are generated.
  - Adding Equipment. This can include anything from investing in the existing equipment to make it faster, to adding production lines, to building a completely new plant.

With both plant and warehouse capacity, the time period for capacity must be considered (annual or long-term models or very seasonal business).

$$\min \sum_{i\in I}\sum_{j\in J} dist_{i,j} d_j Y_{i,j}$$

$$\sum_{i\in I} Y_{i,j} = 1 \ \forall j \in J$$

$$\sum_{i\in I} X_i = P$$

$$\sum_{j\in J} vol_{i,j} Y_{i,j} \le cap_i X_i \ \forall i \in I$$

$$Y_{i,j} \le X_i \ \forall i \in I, \forall j \in J$$

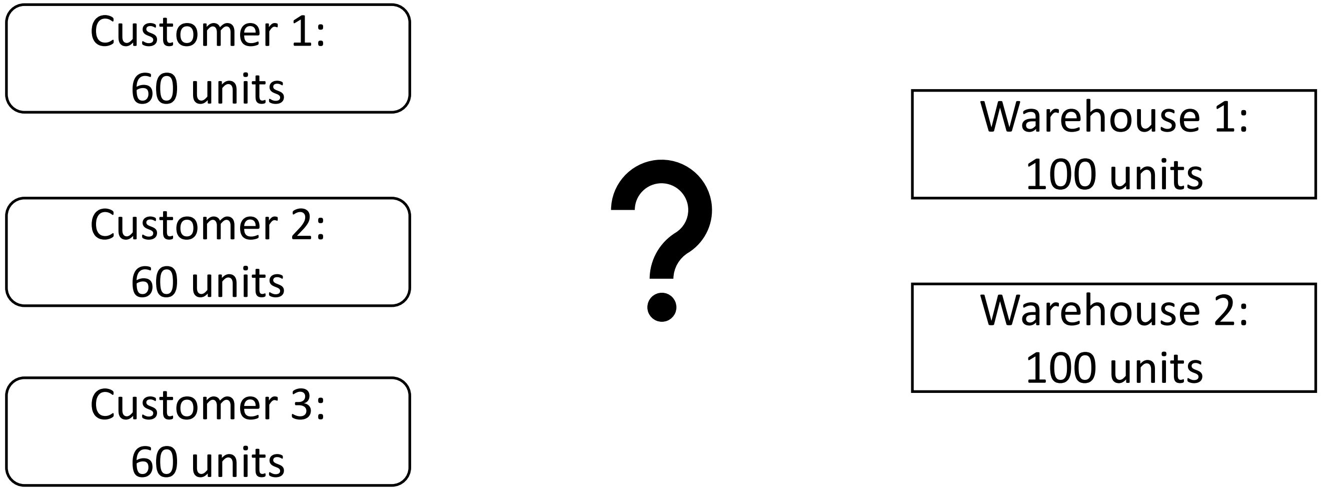
$$Y_{i,j} \in \{0,1\} \forall i \in I, \forall j \in J$$

$$X_i \in \{0,1\} \forall i \in I$$

- cap_i is the capacity of warehouse/plant i:
  - At a warehouse: storage capacity, total throughput capacity, or ability to process inbound or outbound shipments.
  - At a plant: overall time available in regular time and overtime, overall tons or units that can be produced, or maximum amount of a product you can produce.
- $vol_{i,j}$ is:
  - the effective volume of demand for customer j being assigned to warehouse i
  - the amount of a plant's capacity that is consumed by that customer's demand
- It ensures that a warehouse is never assigned more demand than it can handle. If the warehouse capacity is not infinite, this constraint also will ensure that if a customer is assigned to a warehouse, that warehouse must be opened.

It is again possible to create a model such that it is impossible to solve:

- The demand is greater than the capacity constraint
- The constraint on the number of sites conflicts with the capacity constraint
- The capacity required of individual customers is large relative to the overall capacity (capacity is assigned in large chunks), and you are forcing all demand for each customer to be serviced from only one location (single sourcing):



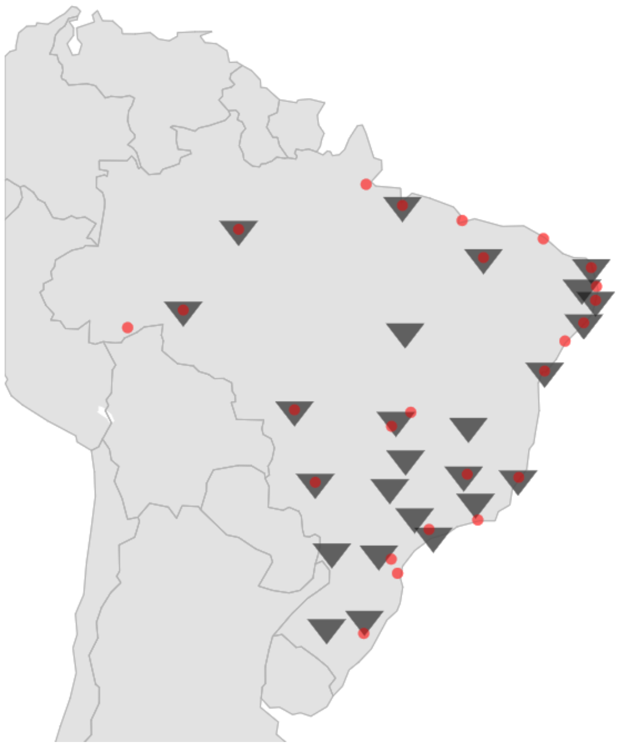
Knapsack problem: it tries to allocate discrete resources in a way that maximizes total value, while staying below a total limit on capacity. It can be hard to solve.

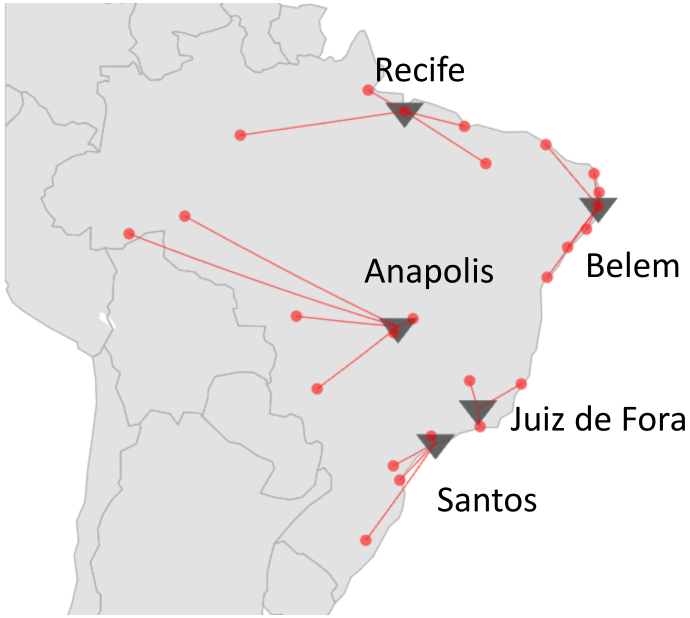
When capacity constraints are set up, a hard knapsack problem could be inadvertently embedded into the network model. If capacity is measured in large chunks, a knapsack problem is embedded into an already difficult problem. This causes three problems:

- The problem is infeasible and in a way that is not clear (the total capacity is more than enough for the demand, but because of the way demand is grouped, it does not all fit into the capacity).
- The solution looks bad. The optimization returns a solution that does not look intuitive. Customers are not assigned to the closest warehouse and may be assigned to warehouses quite far away.
- The runtimes are very long.

Case study: distributor in Brazil
(Watson et al., 2013)

Without capacity constraints:





Warehouse	Throughput
Santos	41545912
Juiz de Fore	23930340
Recife	17936334
Anapolis	7986478
Belem	7265529

The customers along the southeastern coast have the majority of the demand and that placing warehouses close to them is a major driver for minimizing weighted distance traveled for all deliveries.

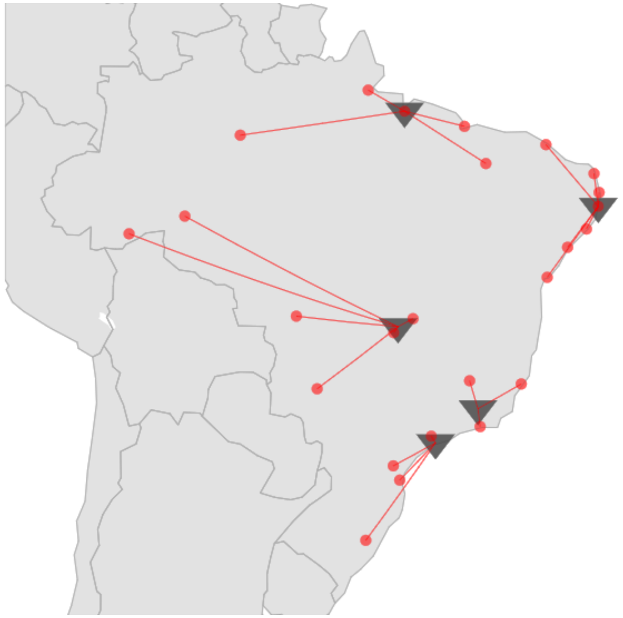
- 25 customer regions
- Each region has a different demand
- The total demand is approximately 100 million units

Model 1: the model is forced to use the same five warehouses selected previously and capacity of each warehouse is set to one-fifth of the total model throughput. No feasible solution exists:

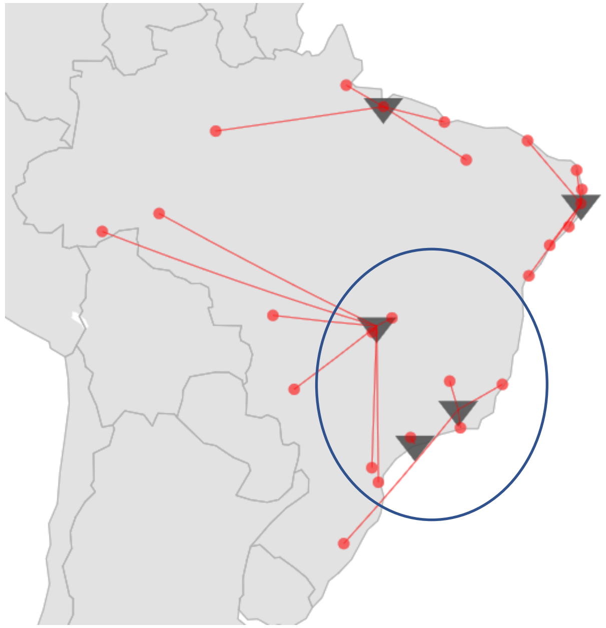
- The 20-million-unit capacity constraint e added at each warehouse attempts to force each warehouse to hold almost exactly that amount with no extra room to use elsewhere as our total flow of goods is approximately 99 million units. Therefore, no warehouse will have space for more than 1 million additional units within the final solution.
- A customer may receive product from only one warehouse location

Therefore, if customer demand cannot be equally split up in 20 million unit buckets, then this problem becomes infeasible.

Model 2: the model is forced to use the same five warehouses selected previously and capacity of each warehouse is set to 30 million units



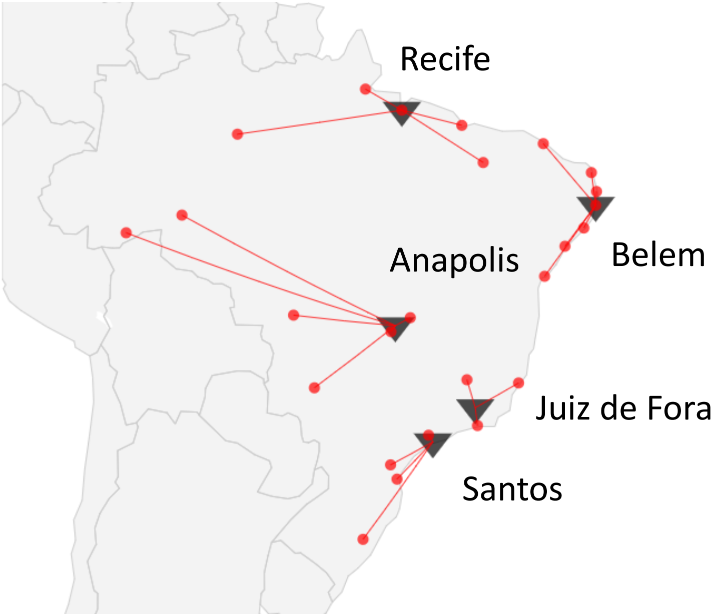
Model without capacity constraints
Average distance to customers: 326km



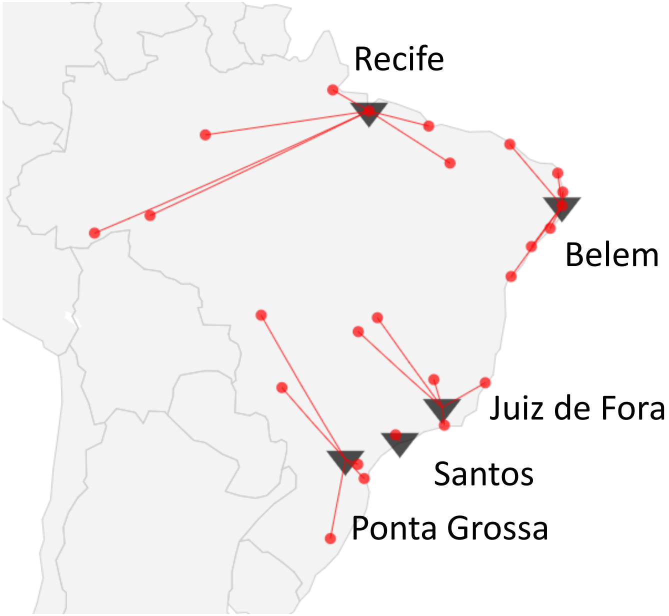
Model 2
Average distance to customers: 412km

Warehouse	Throughput	30 Milinon unit capacity throughput
Santos	41545912	29029226
Juiz de Fore	23930340	28842504
Recife	17936334	17936334
Anapolis	7986478	15591000
Belem	7265529	7264529

Model 3: the model can select any five warehouses and capacity of each warehouse is set to 30 million units



Model without capacity constraints
Average distance to customers: 326km



Model 3
Average distance to customers: 341km

Warehouse	Throughput	Best 5 with capacity throughput
Santos	41545912	29029226
Juiz de Fore	23930340	29042104
Recife	17936334	17936334
Ponta Grossa	-	-
Anapolis	7986478	14712066
Belem	7265529	7946863

- When adding constraints to any model, the solution will never improve. And if the constraint is meaningful enough, the results of the objective may get much worse.
- Although constraints do ensure worse results, they are sometimes a necessity in order to produce implementable results. The key to their use in modeling is to ensure that only meaningful constraints that aren't causing unneeded pressure on the model or, even worse, leading us to a problem with an infeasible solution are applied.
- Capacity constraints don't necessarily change the locations of facilities, but they do have the impact of changing the warehouse to customer assignments.

Outbound transport cost

Transportation is often the most important cost in a network design study.

A change in the number and location of facilities often impacts transportation costs more than any other cost in the supply chain.

- Outbound transport cost: cost to ship outbound from a facility
- Inbound transport cost: cost to get product into the facility, from a plant to a warehouse, before it moves on to a customer

Aim: to locate a given number of facilities, minimizing transport cost, not the average distance.

In many cases, the solution will change when transport costs are added to the model:

- Minimum costs in transport. When a full truckload is shipped, there is often a minimum charge assessed independent of the distance
- Regional differences in transportation costs. Transportation rates are not symmetrical due to the imbalance of supply and demand. By adding these types of transportation costs to the analysis, the optimization may be able to locate facilities in areas with more favorable rates.
- Different customer shipment profiles—In a single network some customers may always order a full truck's worth of products at a time, while others may only order a half a truck. In these cases, the optimization has an incentive to select facility locations closer to the more expensive half-truck quantity customers.

$$\min \sum_{i \in I} \sum_{j \in J} trans_{i,j} d_j Y_{i,j}$$

- Objective: minimize the total transportation costs from the facilities (plants or warehouses) to the customers
- The matrix $trans_{i,j}$ represents the cost to send one unit of demand from facility i to customer j

$$\sum_{i \in I} Y_{i,j} = 1 \ \forall j \in J$$

$$\sum_{i \in I} X_i = P$$

$$Y_{i,j} \leq X_i \ \forall i \in I, \forall j \in J$$

$$Y_{i,j} \in \{0,1\} \forall i \in I, \forall j \in J$$

$$X_i \in \{0,1\} \forall i \in I$$

- The demand is expressed in total (usually annual) and not shipment by shipment.
- The transportation costs are expressed in terms of a cost per unit. That is, we are multiplying transportation costs by units of demand, not by shipments.
- A transportation cost for every possible source and destination is to be filled, not just the source and destination shipments used in the past.

The demand is expressed in total, not shipment by shipment, otherwise:

- The model becomes unwieldy, because if we model every shipment, we have to create a model with potentially hundreds of thousands of shipment points instead of several hundred customer points.
- A more accurate model has not been created, since the model is run for the future years (no reliable forecast can be performed)

Transport cost per unit:

In standard transportation reports or quotes from carriers, rates expressed as a cost per load, cost for the pallet, or cost per package, falling into the same problem mentioned above. Standard transportation rate structures need to be converted into a simple cost per unit. These structures are driven by different modes of transportation:

- Commercial Truckload or Full Truckload (TL or FTL). A firm is hiring an entire truck to drive from one location to another. The firm is charged for the use of the full truck.
    - The trucking company manages where the truck comes from before the delivery or where it goes after.
    - Rates are typically structured either as a flat rate from one point to another or as a cost per mile or kilometer. The rates charged can vary depending on where the shipment starts and ends.
    - If the carrier charges a cost per distance, they may sometimes add a fixed charge on top of that and enforce a minimum charge to prevent very low costs for short moves. The carrier contract may also include a separate fuel surcharge.
    - Each of these trucks will have a capacity typically measured in total weight, total cube, or total number of pallets.

- Private/Dedicated. A private fleet is one in which a firm owns its own trucks. These trucks are used to either deliver full loads to the customer or make multiple stops. A dedicated fleet is similar to a private fleet except that it is owned (and maybe operated) by a third party but dedicated to a single client.
  - The rates for private or dedicated are usually a cost per mile with an optional fixed cost.
  - Truck returned must be managed
  - Often, with a private fleet, firms can use the trucks to make deliveries to their customers and then also pick up product from their nearby suppliers on their way back to the origin (a roundtrip or backhaul).
- Less-than-Truckload (LTL). In the TL market, a truck comes to a firm's dock door, picks up the product, and directly delivers it to the destination.
  - A firm goes to the LTL market when it has a small load to ship to a destination where it is not economical to hire an entire truck to make the delivery. In this market, the LTL trucking company picks up the firm's small load on a route with other firms' pickups and deliveries. This load then rides with other loads through multiple hubs, where it is placed on different trucks with loads having similar geographic destinations, until eventually loads are placed on a truck departing a final hub to make the trip to complete a route of deliveries to customer locations.
  - A firm does not need to worry about these intermediate activities completed by the LTL carrier.
  - The LTL trucking company quotes a single price from origin to destination even if they route the product through different hubs each time. There are break points in the rates based on some form of capacity level (the more product shipped, the lower the cost per unit to deliver it). Minimums and fuel surcharges are often assessed.

- Parcel. The parcel mode is a lot like LTL but with even smaller shipments sizes.
  - A firm is typically charged based on the source, destination and package weight.
  - For heavier packages or shipments, this can be quite an expensive transport mode. However, for small packages, it can actually be much cheaper than shipping via LTL.
- Ocean. Ocean transport is similar to truckload transport but a firm’s product moves in containers across the ocean between port cities.
  - Ocean rates are usually quoted as the cost of a single container from one location to another.
  - The costs can be from port to port or from door to door (where the truck or rail transport to the first port and from the last port is included in the rate).
- Rail. Rail transport moves a firm’s product from one railhead to another.
  - Rail rates are typically expressed in terms of the cost of a rail car to get from one point to another.
  - Like ocean transport, these rates can be between two railheads or from door to door.
- Intermodal.
  - These rates typically are expressed as door-to-door rates because the shipping firm handles the full move from origin to final destination.
  - This type of transport can be very economical for cross-continent moves.

- Multistop. The multistop mode is like the truckload moves, but involves several stops.
    - A firm has several shipments to deliver. The shipments are not all large enough to use a full TL for each however. But together they are too big for an LTL shipment as well. In multistop transit a truck picks up a firm's load at one facility and then makes deliveries to several destinations. Unlike LTL, the truck is dedicated to just this firm's loads.
    - The rates are usually expressed in terms of a cost per mile or kilometer with a stop-off charge. The stop-off charges may be waived for the first stop and commonly increase with the number of stops being made on a single route.
    - Often, a trucking company will actually put a limit on the number of stops a single truck will make. There also may be restrictions on the "out of route" miles a single truck may travel within a single route. Out of route is a measure of how far the route deviates from a direct route from the source to the final destination.
    - The multistop mode cost falls between TL and LTL above.

The more is shipped at one time, the lower the transportation costs. That is, if the same total amount is shipped, the LTL cost will be lower than parcel, the TL lower than LTL, and Rail lower than TL.

Basic formula to convert the costs to a cost per unit:

$$trans_{i,j} = \frac{load_{i,j}}{avg_{i,j}}$$

- $trans_{i,j}$ is the cost to go from point i to point j to satisfy all the demand at point j that is satisfied from shipments from point i
- $load_{i,j}$ is the cost for the load to move from point i to j
- $avg_{i,j}$ is the average size of the load. To estimate $avg_{i,j}$ (units of measure are in pounds):
  - For parcel and LTL, the cost per load is dependent on the size of the shipment.
  - For TL (and rail, intermodal, ocean, private/dedicated—which are calculated like TL), the average size is the amount that is actually shipped.

Estimation of the TL load cost:

Origin	Destination	Shipment Wt (lbs)	Distance (miles)	TL $/mile	Min charge	Distance related cost	TL load cost	Transport cost per lb
Chicago	Dallas	38000	967	1.80	400	1741	1741	0.046
Chicago	Indianapolis	27000	183	1.8	400	329	400	0.015
Atlanta	Miami	40000	663	2.00	375	1326	1326	0.033
Atlanta	Chattancooga	30000	118	2.00	375	236	375	0.013

Estimation of the average shipment size:

Origin	Destination	Total truckloads	Total Wt shipped (lbs)	Avg. TL Wt (lbs)
Tulsa	Houston	26	832858	32033
Tulsa	Albuqurque	35	1246175	35605
Tulsa	St.Louis	54	1084508	38602

Potential exercise in the exam

A rate from every source to every destination is needed. However, this information could not be available, because:

- New lanes are to be modelled
- Negotiated rates available for existing lanes may not be applied to new lanes
- Characteristics of available data are not suitable
- There is no time to gather data

The shipping cost need to be estimated, by:

- Using a simple cost per mile or per kilometer based on the average rates available today
- Using nonnegotiated rates and extrapolating. These rates will most likely be higher than what is paid today. Therefore the output lane costs for the current lanes can be taken and reduced (or increased) by a uniform percentage so that the total cost on the lanes matches the total cost from historical data. Then this blanket percentage reduction (or increase) can be applied to all the rates in new data for new scenarios.
- Using high-level industry benchmarks, that is industry-wide average cost per ton- miles published by various independent sources. These rates are commonly expressed as the average cost to move one ton of an item one mile.
- Using regression analysis. A simple regression plots the distance of the shipment by its cost, then a regression formula to determine a cost based on how far the shipment goes can be developed.

Example of calculating an average rate for a sample set of data:

Shipment number	Origin	Destination	Shipment Wt	Freight cost ($)	$/mile	Avg $/mile	Std Dev $/mile	Coeff. Of Variation
1	Atlanta	Toledo	39800	1286	1.94	1.91	0.05	3%
2	Atlanta	Columbus	33503	1043	1.84			
3	Atlanta	Lima	38138	1149	1.96			
4	Atlanta	Youngstown	30500	1376	1.86			
5	Atlanta	Canton	34313	890	1.93			
6	Atlanta	Raleigh	32175	625	1.59	1.68	0.08	5%
7	Atlanta	Charlotte	37409	417	1.70			
8	Atlanta	Wilmington	34305	730	1.76			
9	Atlanta	Elizabeth City	36695	996	1.73			
10	Atlanta	Rocky Mount	39253	742	1.60			
					Aggregate statistics	1.79	0.14	8%

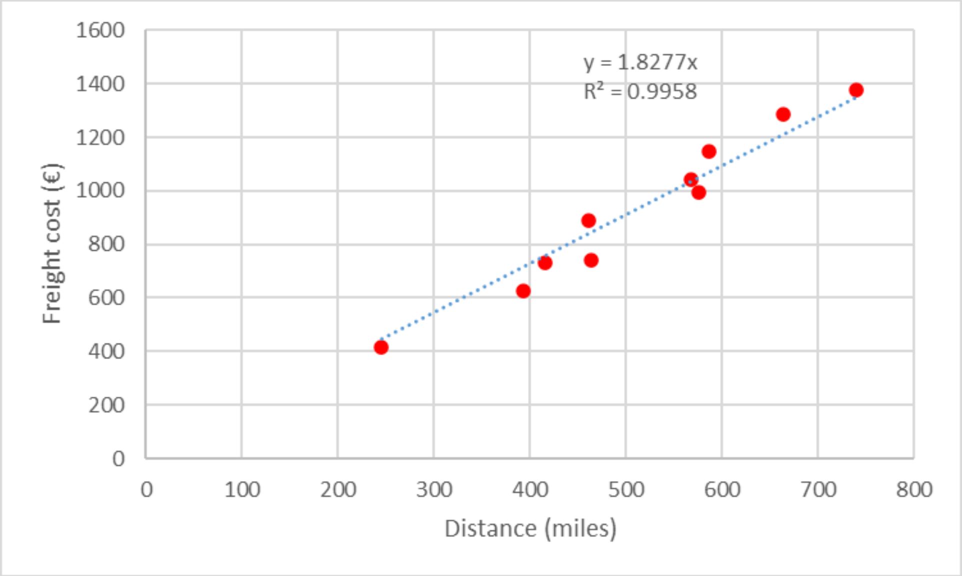
To test whether the average is a good measure, the standard deviation and the coefficient of variation (Std Dev / Average) are estimated. The standard deviation gives an absolute measure of variability and the coefficient of variation normalizes the measures and expresses it as a percentage.

Regression can be valuable for helping determine a good cost to use for sources and destinations where no current shipment activity or rates are available.

Shipment number	Origin	Destination	Shipment Wt	Freight cost ($)
1	Atlanta	Toledo	39800	1286
2	Atlanta	Columbus	33503	1043
3	Atlanta	Lima	38138	1149
4	Atlanta	Youngstown	30500	1376
5	Atlanta	Canton	34313	890
6	Atlanta	Raleigh	32175	625
7	Atlanta	Charlotte	37409	417
8	Atlanta	Wilmington	34305	730
9	Atlanta	Elizabeth City	36695	996
10	Atlanta	Rocky Mount	39253	742

Potential exercise in the exam

- The distance for each shipment (or lane) is calculated.
- The shipments are plotted on a chart with freight costs and distance
- A regression analysis is performed (using capabilities within Excel) by considering the entire set of lanes



Potential exercise in the exam

A more detailed analysis can be accomplished by extending the same type of analysis across multiple origins and destinations to create a rate matrix, showing cost per mile rates on a more detailed spatial level.

With multistop, the final warehouse may send out a truck that makes several stops at different stores on the same route. So the transportation costs are no longer a simple point-to-point rate. The transportation rates will depend on which stores are on which route.

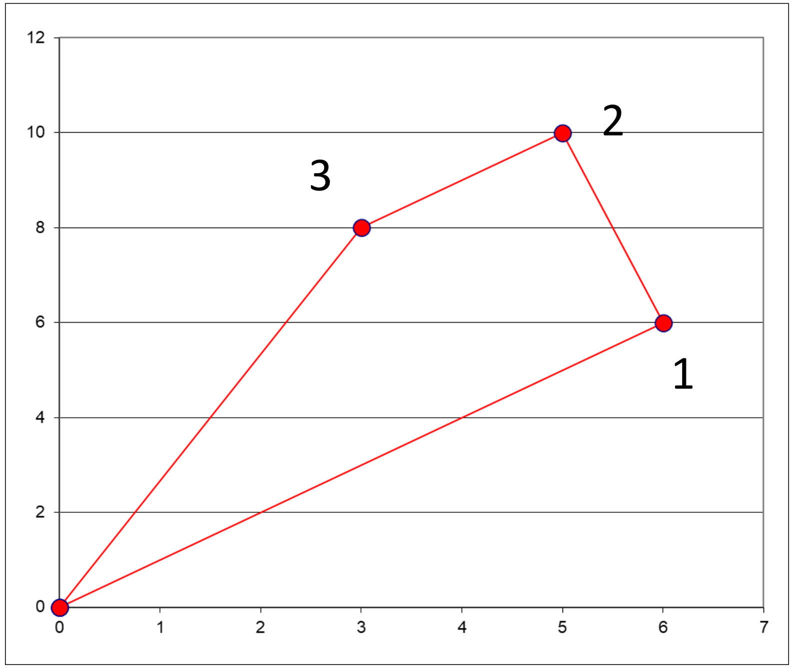
Basic principles of a robust multistop approach:

- Worst case: the costs of multistop routes are somewhere between the cost of TL moves and those of LTL moves.
- A scaling factor, obtained from the comparison of the equivalent costs of the point-to-point rate with the cost of the route, can be used to adjust the costs.

Examples:

A multistop route from a warehouse to three stores, with straight-line travel distance and costs that are directly proportional to distance.

The warehouse is at point (0,0) and the trucks must all start and end here.

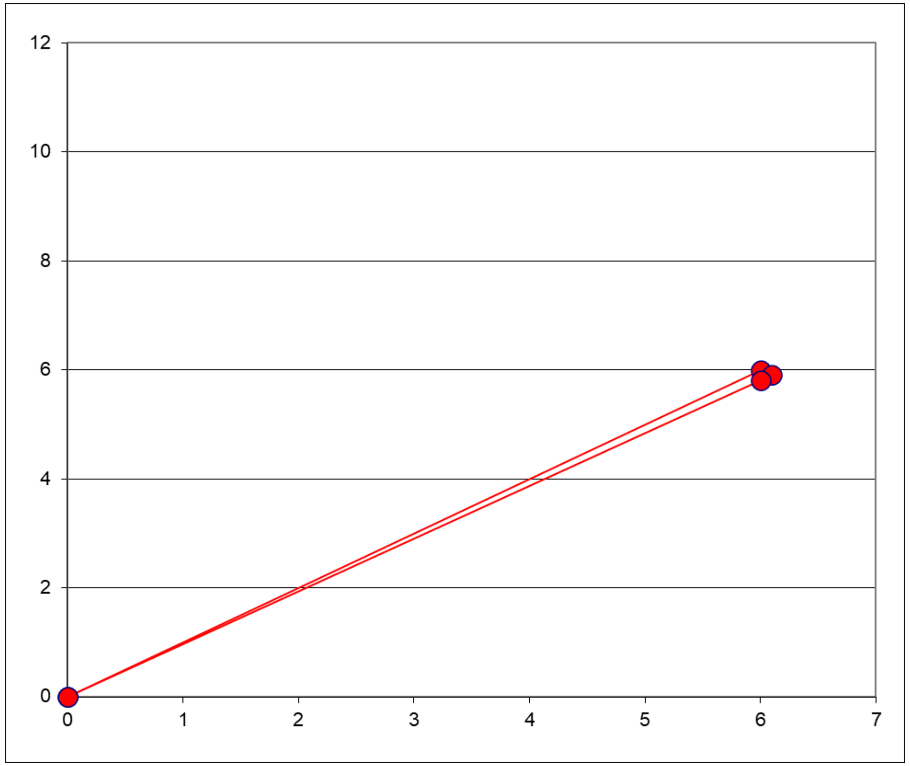


Distance between two points:

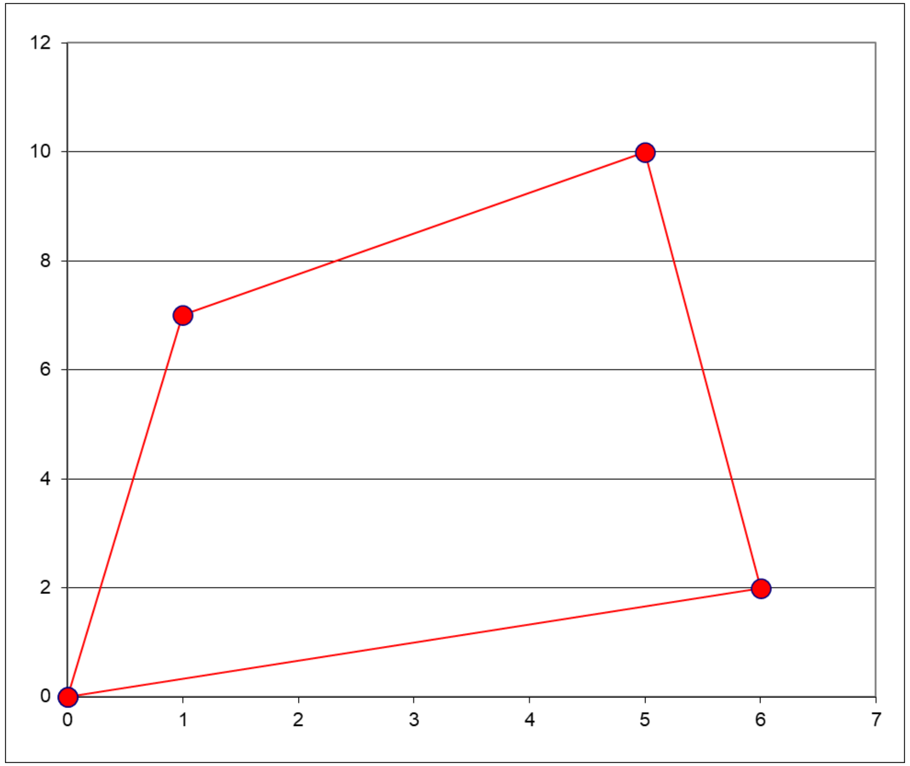
$$dist = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

- The total distance of the route is 23.98
- In the network design formulation however, the distance from the warehouse directly to each of these points separately is considered:
  - The formula to calculate the distance from the depot to each customer.
  - Each value is multiplied by two for the trip back
  - The three distances are added together.
  - This gives a total distance of 56.42
- To make a fair comparison of the multistop route to the 56.42:
  - A multistop route is adopted to perform more frequent deliveries. So, in this case, each route delivers one-third of the total requirements for each customer. Therefore, running the multistop route three times allows to make a fair comparison to the direct full truckload delivery distances (56.42). Running the route three times gives a total distance of 71.94.
  - Full requirement could be delivered to each store with a single truckload shipment (total distance of 56.42). Or three deliveries on multistop routes can be made (total distance of 71.94).
  - It makes sense that the route distance is longer for the multistop in this case (the benefit of more frequent deliveries comes at an additional cost).
- To approximate multistop costs using TL rates within the model, the TL rates are multiplied by 1.28 (71.94/55.42).

The factor 1.28 will change based on the structure of the routes as well. If the stores are all far away from the warehouse, but close to each other, the factor may be much smaller.

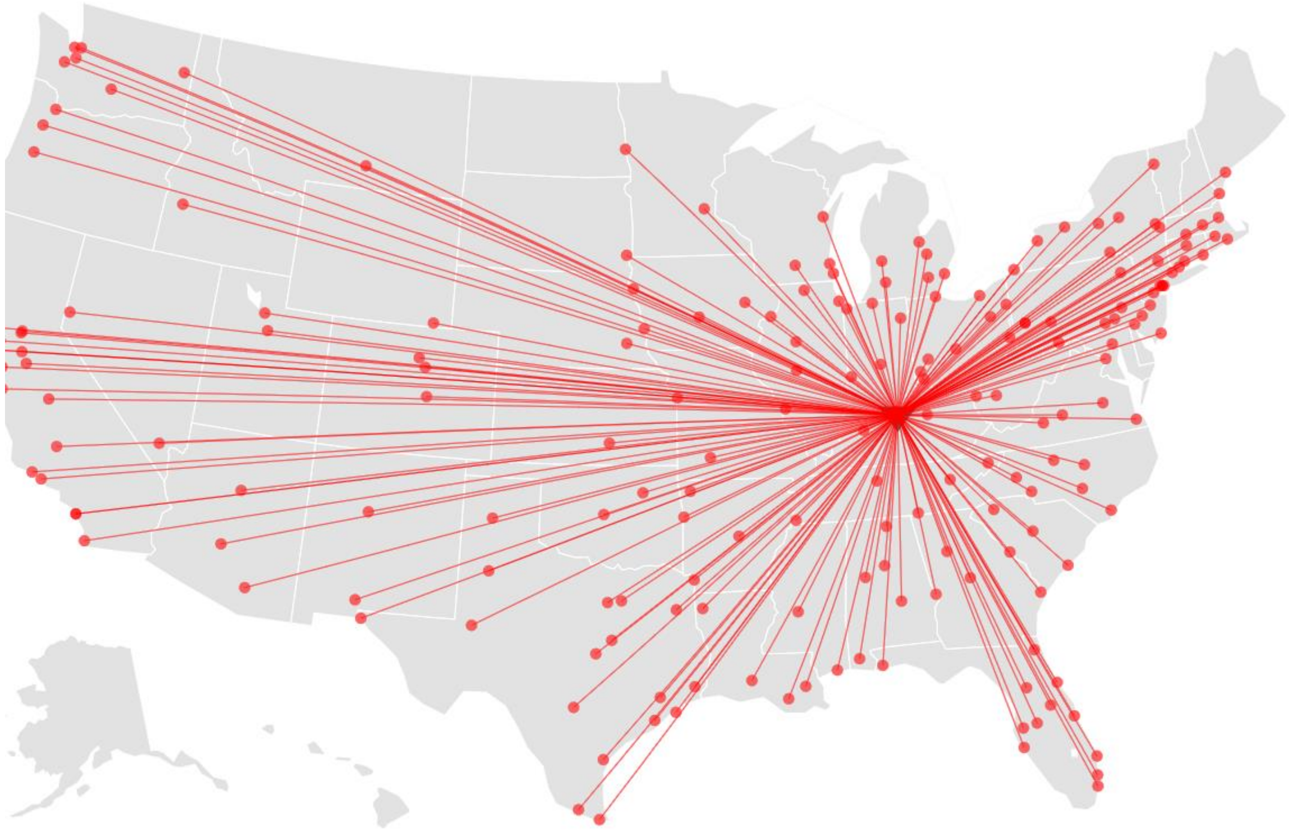


- the total two-way distance is 50.6
- running the route three times is 51.3
- the factor is 1.01



- the total two-way distance is 49.2
- running the route three times is 79.4
- the factor is 1.64

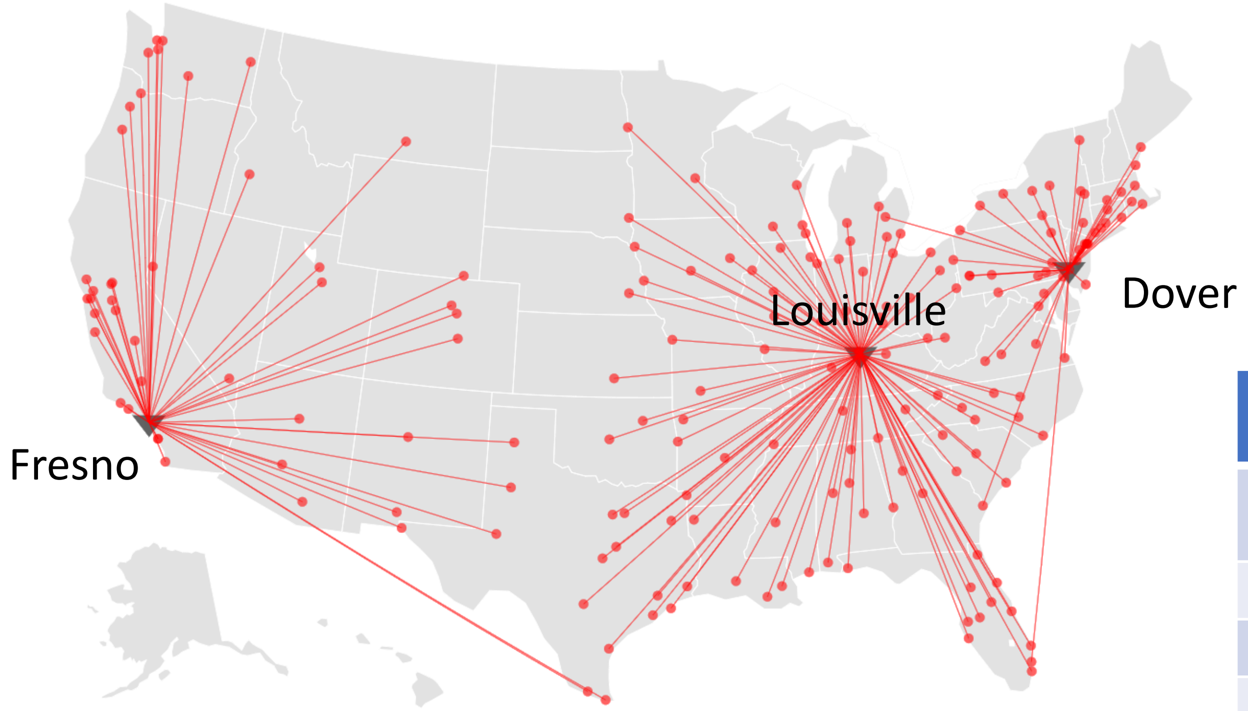
Case study: distributor in the U.S.
(Watson et al., 2013)



- The current cost of this supply chain is $16.6 million
- The average distance to customers is approximately 975 miles

Distance band (miles)	Total packages	% of total	Cumulative %
300	259000	8	8
600	919000	28	36
900	949000	28	65
5000	1135000	35	100

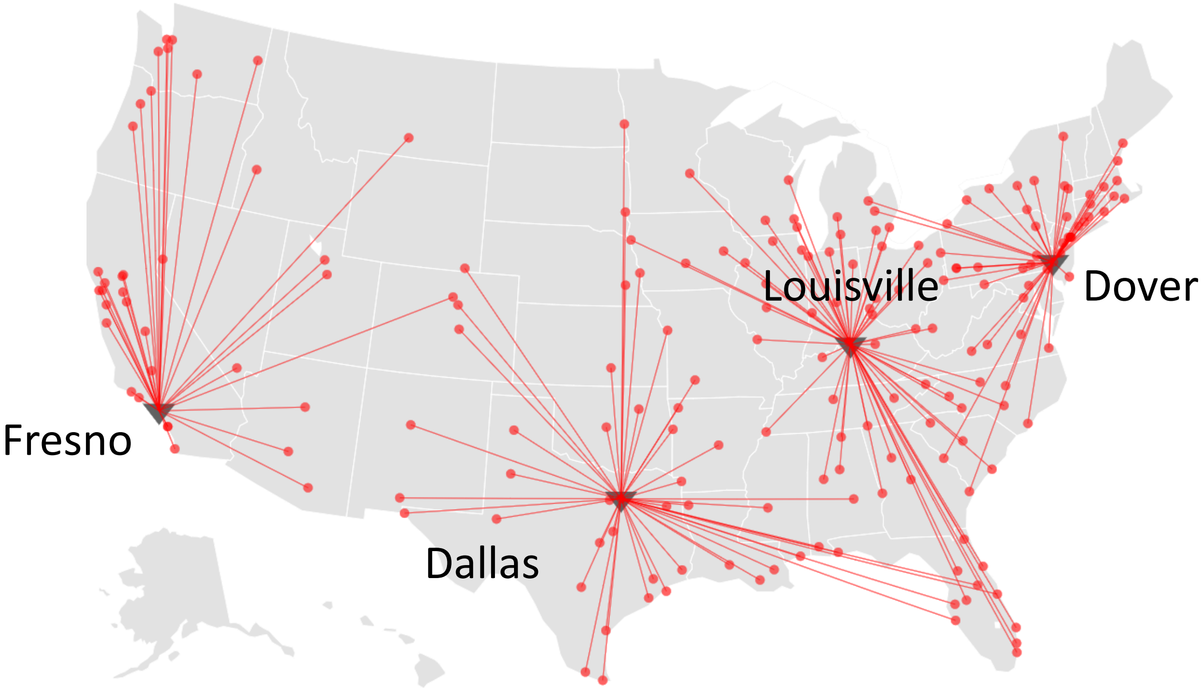
This firm wants to determine whether they can reduce costs by opening 2 additional warehouses.



- The total cost decreases to $15.2 million
- The average distance dropped to approximately 430 miles

Distance band (miles)	Total packages		% of total	
	Current network	3 warehouses	Current network	3 warehouses
300	259000	1347000	8	41
600	919000	1088000	28	33
900	949000	476000	29	15
5000	1135000	351000	35	11

This firm now wants to understand the marginal value of a fourth warehouse. They determined that if they could get somewhere close to another $1.4 million in savings, it might be worth serious consideration.



- The costs decrease to about $15.0 million, which is only a $200,000 savings from their previous 3 warehouse solution (about a 1% reduction).
- This $200,000 difference indicates the threshold for the fixed cost (if the fixed cost of the additional facility is less than $200,000, it would still be cost beneficial to open the fourth facility).

Fixed and variable facility cost

- Fixed costs are costs that are incurred each year (or month) independent of how much volume is handled at the facility (e.g. building the site; expanding the site; adding equipment to the site like extra lines or equipment in plants, additional racking or automation in warehouses; paying taxes; or staffing the location)
- Variable costs are those that depend on the actual number of units that are made at a facility or that pass through a facility.

- If all the costs at a facility are modelled as fixed, then the optimization has the incentive to minimize the number of facilities it opens and, when a facility is open, it will focus on maximizing the amount of product put through that facility.
- If all the costs at a facility are modelled as variable, then the optimization has an incentive to use the facilities with the lowest variable cost and no incentive to avoid adding additional facilities.

- The fixed costs can be considered as a step function (several fixed costs to add, depending on how much product flows through or is made at the facility).
- The variable costs depend on the size of the facility, or the economies of scale. That is, as a facility gets larger, the cost per unit may go down. This can be captured with the step sizes with a lower variable cost for each step size.

Facility location planning

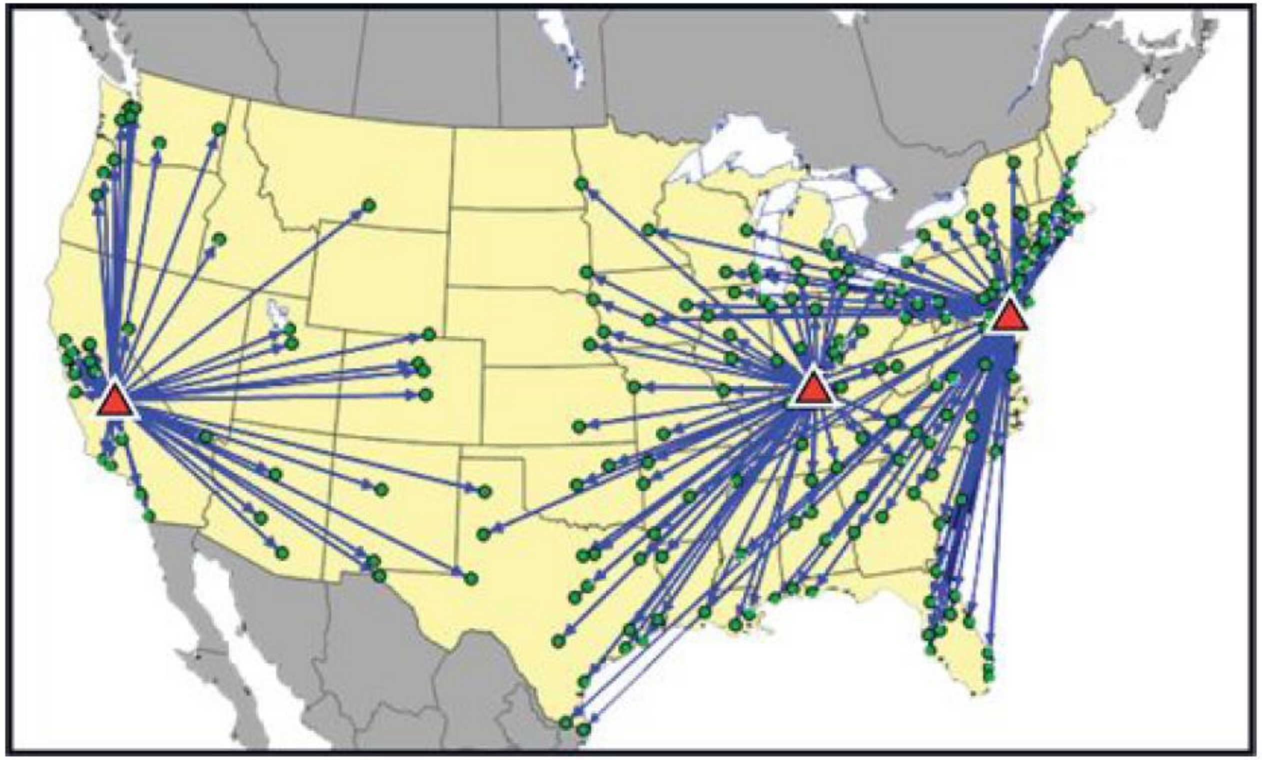
$$\min \sum_{i\in I}\sum_{j\in J}(trans_{i,j}+facVar_i)d_jY_{i,j}+\sum_{i\in I}\sum_{w\in W} facFix_{i,w}X_{i,w}$$

1. $\sum_{i\in I} Y_{i,j} = 1\ \forall j\in J$
2. $\sum_{i\in I}\sum_{w\in W} X_{i,w} \geq P_{min}$
3. $\sum_{i\in I}\sum_{w\in W} X_{i,w} \leq P_{max}$
4. $\sum_{w\in W} X_{i,w} \leq 1\ \forall i\in I$
5. $\sum_{j\in J} vol_{i,j}Y_{i,j} \leq \sum_{w\in W} cap_{i,w}X_{i,w}\ \ \forall i\in I$
6. $Y_{i,j} \leq \sum_{w\in W} X_{i,w} = 1 \forall i\in I, \forall j\in J$
7. $Y_{i,j} \in \{0,1\} \forall i\in I, \forall j\in J$
8. $X_{i,w} \in \{0,1\} \forall i\in I, \forall w\in W$

- The set W represents the set of facility options. Each option represents a different sizing decision for the same physical location (e.g. W represent the choice of building a small, medium, or large warehouse at each potential site). This set of options allows to model changes in fixed and variable costs for each option.
- The decision on whether to open a facility has been generalized to include the choice of facility option. Variable $X_{i,w}$ will be set to 1 if and only if facility i is opened with option w.
- Constraint 4 ensures that more than one facility option for each facility that is opened is not selected.
- Facility fixed (ʄacFix) and variable costs (ʄacVar) are added to the objective function. So the optimization will consider all the costs. The optimization will want to incur a fixed cost only if it absolutely needs the capacity or if the savings from the variable or transportation costs make it worthwhile.
- Constraint 5 expresses limits to the capacity of a facility.
- The model allows the solver engine to determine how many facilities to open. Constraints 2 and 3 restrict the total number of opened facilities to be between P_{min} and P_{max} (without facility fixed costs, there would have been no disincentive for the model to simply open all facilities).

Case study: distributor in the U.S.
(Watson et al., 2013)
Variable costs to the model is added:

- Louisville facility: $2.75 per package
- New locations: $2.50 per package



- The three locations are the, but the territories have shifted
- The objective function considers both the transportation costs and the variable costs. It does not consider distance to the customer. So when the variable cost is different between locations, the decision can change and it can impact the model.
- The model will move more volume to the low-cost warehouses. When the optimization decides to move more volume to these low-cost locations, transportation costs will rise, but the overall result is a lower cost.

To estimate variable costs the following elements should be considered:

- Labor Costs
- Utility Costs. The more product that moves through a facility, the more resources (e.g. electricity and water) the facility uses
- Material Costs

It is easier and cleaner to break out the costs by major activity to help quantify the costs (e.g. putting product away, for storing the product, and for picking and shipping the product to customers).

When there are many different sites and complexity in breaking out the variable costs, a regression analysis can be used:

- The total cost of the site is the dependent variable
- The total throughput of the site is the independent variable. Other independent variables can be added (e.g. product type or type of facility)
- The slope of the line provides information on the variable cost
- The y-intercept of the regression can provide insight into the overall fixed cost per facility

To estimate fixed cost:

- When a scenario is run with one more facility allowed than another scenario, the difference in cost is the fixed cost threshold (e.g. if a model comes back with a four-facility solution at a cost of $10.8 million and the five facility solution is $9.1 million, then the cost of the extra warehouse needs to be less than $1.7 million)

Reasons to include fixed costs:

- The fixed costs vary significantly from location to location
- Capacity is to be considered and the shifts are best treated as fixed costs
- Capital investment decisions are to be addressed to expand or improve existing locations or open new sites (investment is to be put in the same time frame as other costs)

Information on the fixed and variable costs are needed, however the accounting systems containing financial data may just report on total costs:

- Consider a firm’s detailed list of expense categories for a facility (expenses that the accounting team tracks and reports on every month)
- The cost of each item in the list is categorized into the percent fixed and percent variable

Three-echelon supply chain

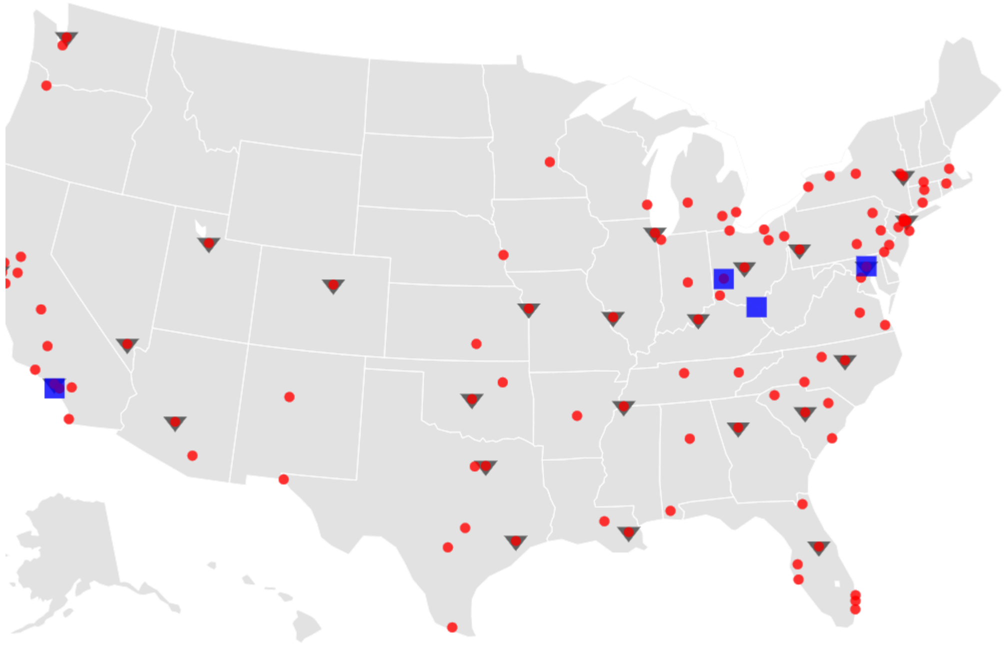
Up to this point, only a two-echelon supply chain has been considered (a facility and the customers it services).

Examples:

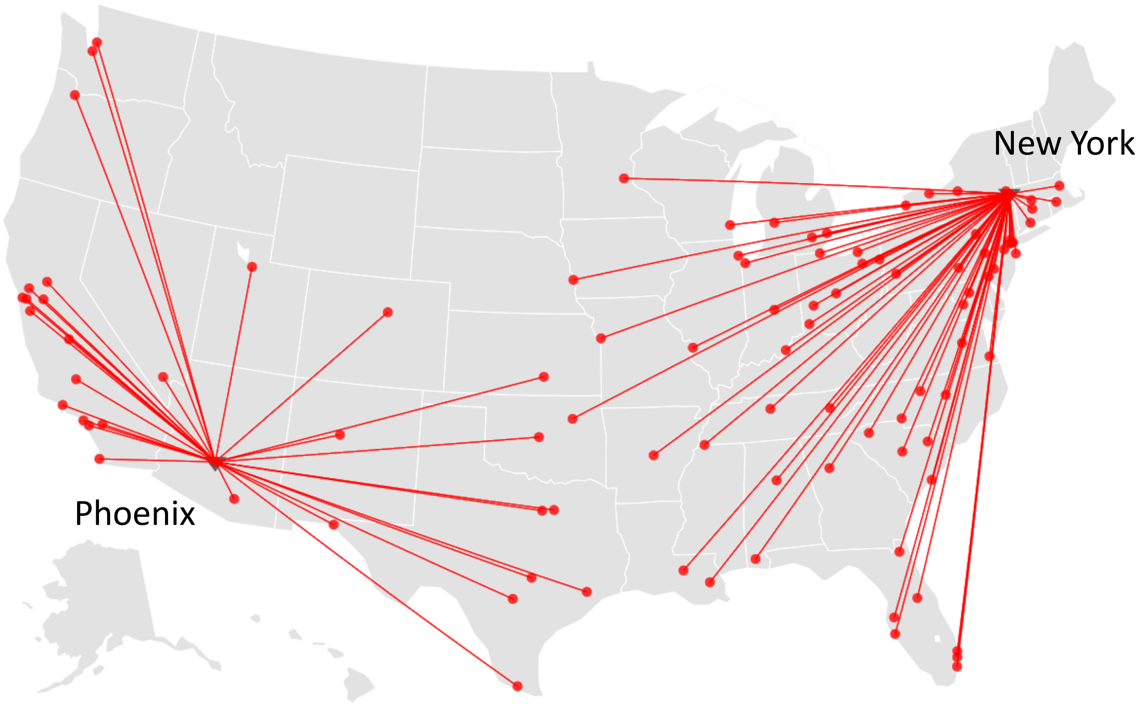
- A set of plants or suppliers that ship to warehouses, and then the warehouses, in turn, ship to customers.
- A supply chain in which a group of raw material suppliers ships to a plant, and the plant then ships to the customers.

Case study: network in the U.S.

(Watson et al., 2013)



- Each plant (marked as blue rectangle) specializes in making just one of the four product families.
- The plants and warehouses are set up to handle rail shipments. Therefore, shipments from the plants to the warehouses are a mix of both truck and rail transit. Outbound to customers, product travels in mostly full (but not always) truckload shipments.
- The shipments from a plant to a warehouse cost $0.07 per ton-mile
- The shipments to a customer cost $0.12 per ton-mile
- In both cases, the rate structure has a minimum charge of $10 per ton



Aim: defining a new network structure that reduces the current supply chain costs

In the current supply chain:

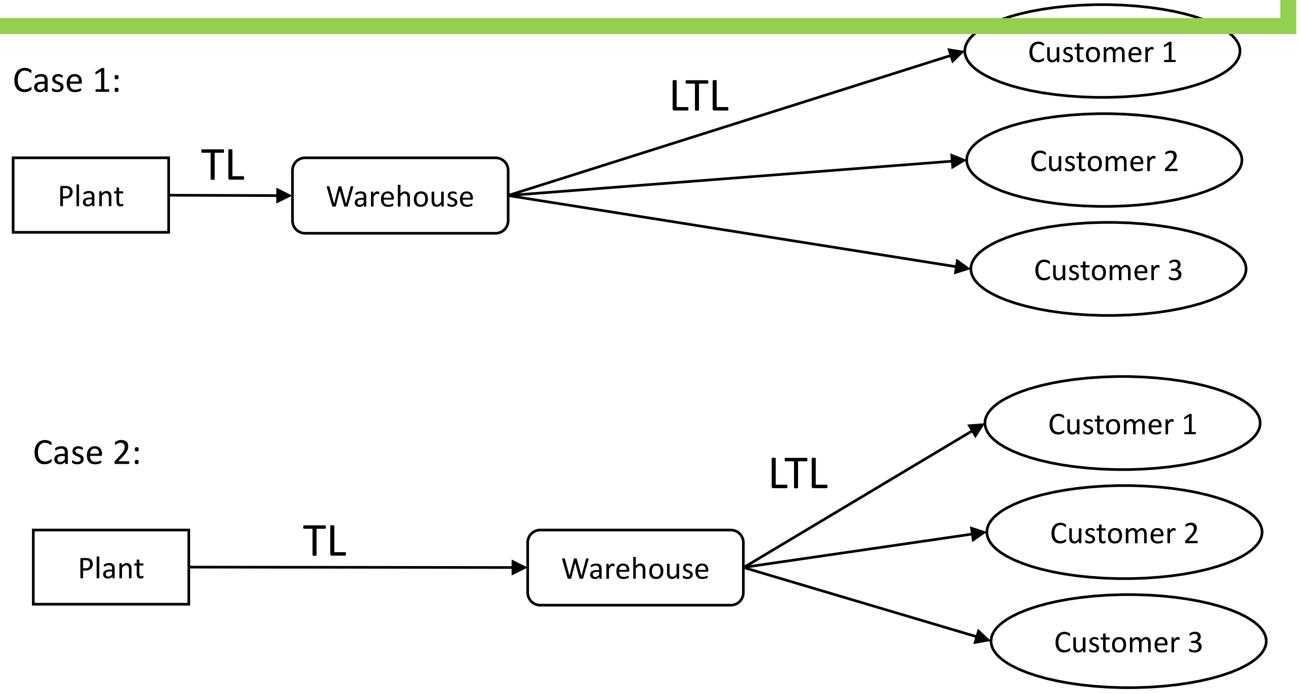
- The market is served from two warehouses
- Transportation cost is $254 million per year:
    - $133 million is the cost to get product from the plants to the warehouses
    - $121 million is the cost to get product to the customers
- Only 21% of the demand lies within 200 miles of their current warehouse locations and only 32% of demand is within 400 miles

Locating warehouses given fixed locations of plants and customers to minimize cost:

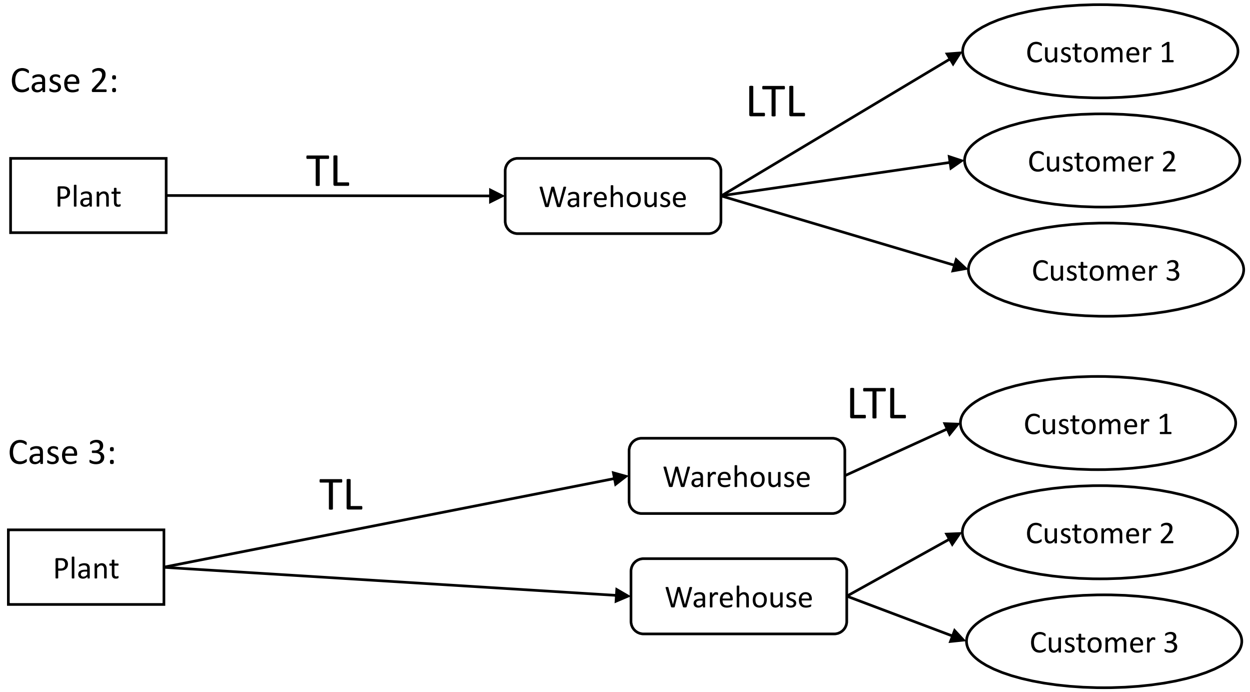
- The plants ship to the warehouses and then the warehouses ship to the customers
- Unlike previous models, the location of the existing plants or supply points will have an impact on the optimal location of the facilities as well. Previous model tended to locate facilities close to customers to reduce transportation costs. Now, both the customer demand and the supply points pull on the location of the facility. The aim is to minimize the cost of shipments both to and from the facility.
- The relative strength of these two forces depends greatly on the relative cost difference between the inbound and outbound costs.

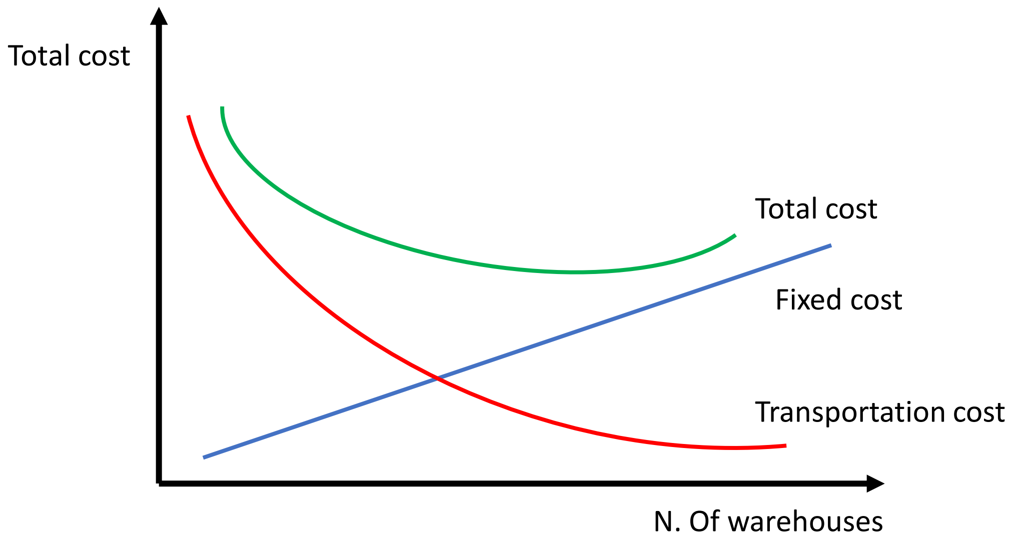
Example:

- Truckload (TL) shipments from a plant to the warehouse.
- The warehouse ships less-than-truckload (LTL) to customer points.
- In Case 1, the plant ships via TL a short distance to the warehouse. In Case 2 the warehouse is much closer to the customers.
- Case 2 have the lower cost (the total cost is minimized by maximizing the distance on TL and minimizing the distance on LTL)



- In Case 3 a second warehouse is added
- The customers still get the same sized shipments and the plant has enough volume that it can ship in full truckloads
- Case #3 has lower transportation costs.
- But which supply chain has the lower cost is depends on the cost of adding the second warehouse
- If the transportation savings offsets the warehouse cost, Case 3 has the lower cost; if not, Case 2 is better.





- The transportation costs decrease as additional warehouses are added.
- These additional warehouses also enable to get closer to customers and reduce transporation costs.
- However, as warehouses are added, the total fixed cost of the warehouses rises.
- The best cost solution is at the bottom point of the total cost curve.

This is a fairly common curve, however each case will be different. In the example, being closer to the customers helped drive down the transportation cost but this is not always the case:

- A plant may require a good deal of heavy raw materials for use in the production process. The resultant finished good sent out to customers is relatively small and light in comparison. So the inbound cost may be expensive in relation to the cheaper outbound transportation.
- A retailer may ship full truckloads of a large assortment of its products to stores but source each product from a separate supplier requiring a lot of small shipments into its warehouse.

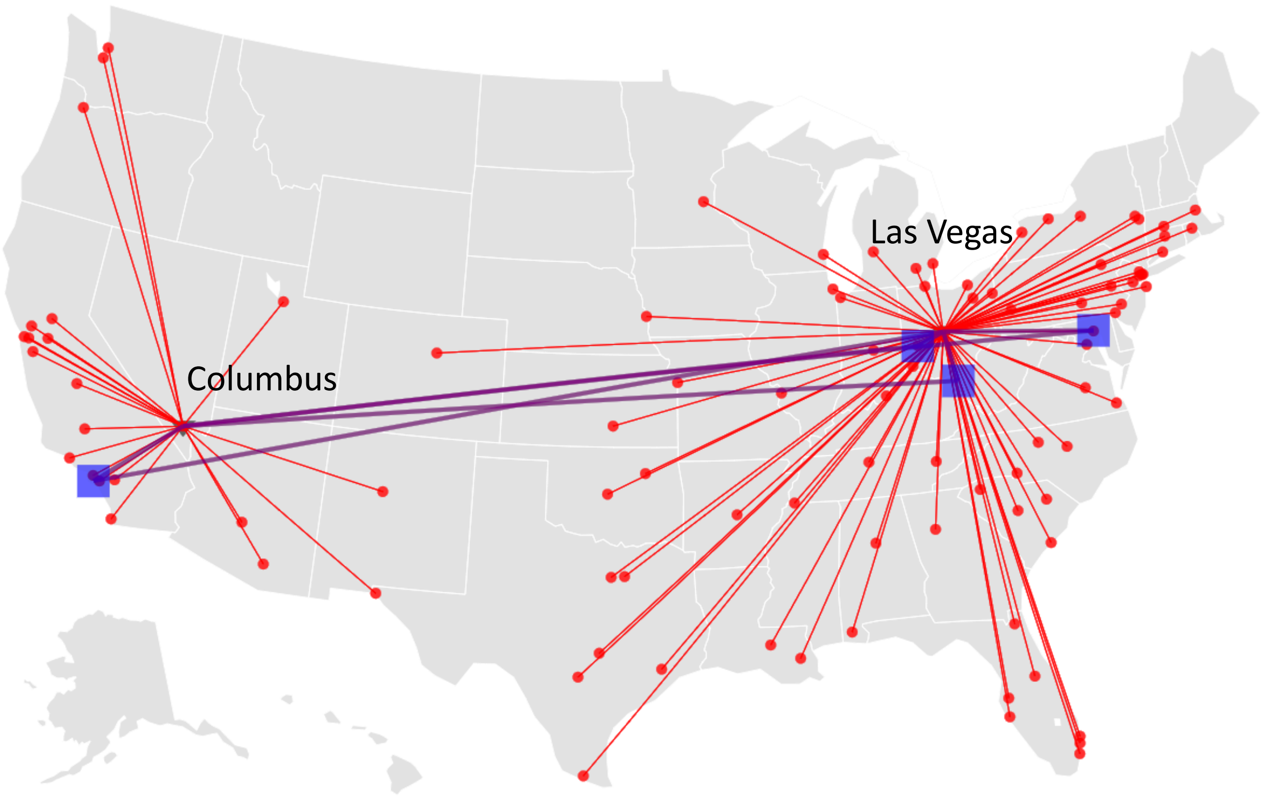
$$\min \sum_{l \in L} \sum_{i \in I} (transPW_{l,i} + pVar_l) Z_{l,i} + \sum_{i \in I} \sum_{j \in J} \left(transWC_{i,j} + whVar_i \right) d_j Y_{i,j} + \sum_{i \in I} \sum_{w \in W} whFix_{i,w} X_{i,w}$$

1. $\sum_{i \in I} Y_{i,j} = 1 \; \forall j \in J$
2. $\sum_{i \in I} \sum_{w \in W} X_{i,w} \geq P_{min}$
3. $\sum_{i \in I} \sum_{w \in W} X_{i,w} \leq P_{max}$
4. $\sum_{w \in W} X_{i,w} \leq 1 \; \forall i \in I$
5. $\sum_{j \in J} vol_{i,j} Y_{i,j} \leq \sum_{w \in W} cap_{i,w} X_{i,w} \;\; \forall i \in I$
6. $\sum_{l \in L} Z_{l,i} = \sum_{j \in J} d_j Y_{i,j} \; \forall i \in I$
7. $\sum_{l \in L} Z_{l,i} \leq pCap_l \; \forall l \in L$
8. $Y_{i,j} \leq \sum_{w \in W} X_{i,w} = 1 \forall i \in I, \forall j \in J$
9. $Y_{i,j} \in \{0,1\} \forall i \in I, \forall j \in J$
10. $X_{i,w} \in \{0,1\} \forall i \in I, \forall w \in W$
11. $Z_{l,i} \geq 0 \; \forall l \in L, \forall i \in I$

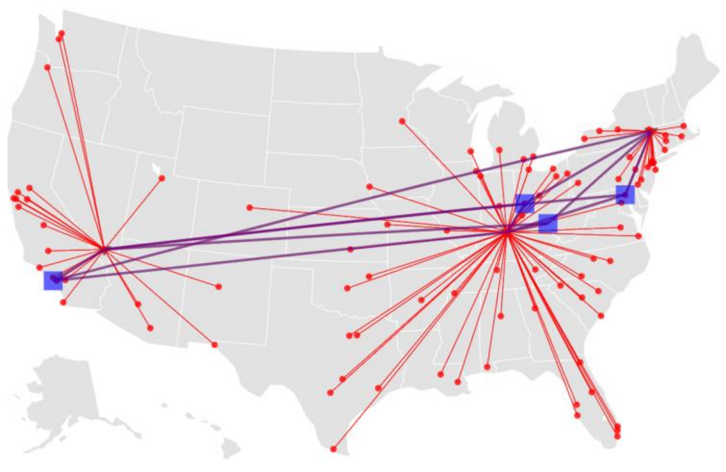
- whVar is the warehouse variable costs, pVar is the plant variable costs, transWC is the transportation cost from the warehouse to customer, transPW is the transportation cost from the plant to warehouse, whCap is the warehouse capacity, pCap is the plant capacity and whFix is the warehouse fixed cost
- The objective function here divides into three distinct sections. The first section computes the total cost of producing or acquiring goods, and shipping them to the warehouse. The second section computes the total cost of handling goods at the warehouses, and shipping them to the customers. The third section computes the total fixed cost of opening the warehouses (so long as the correct total is computed, the objective function can be organized in whatever manner seems more readable and intuitive)
- The Z variable is the flow from the plants to the warehouses.
- Equation (6) represents the “Conservation of flow” constraints that transmit the obligation to ship and produce from the customers to the plants or suppliers). The left hand side sums the inbound shipments over all the plant-to-warehouse lanes that terminate at this warehouse, and the right hand side sums the outbound shipments over all the warehouse-to- customer lanes that originate from this warehouse. By setting these two sums to be equal to each other, warehouses can neither consume nor produce goods, but merely act as conduits to enable more efficient shipping options.

In the case study, transportation costs were by far the dominant cost in the warehouse network, so the function is set to minimize that specific cost area only.

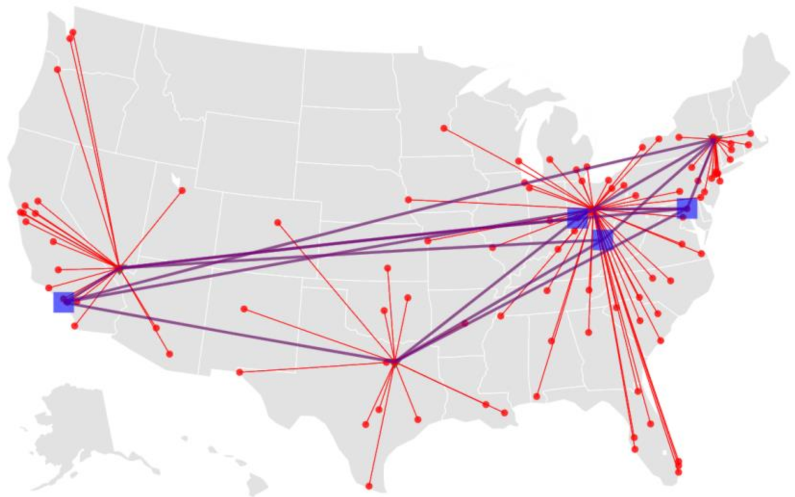
The trade-off is between the inbound cost and the slightly more expensive cost to transport product outbound to the customers.



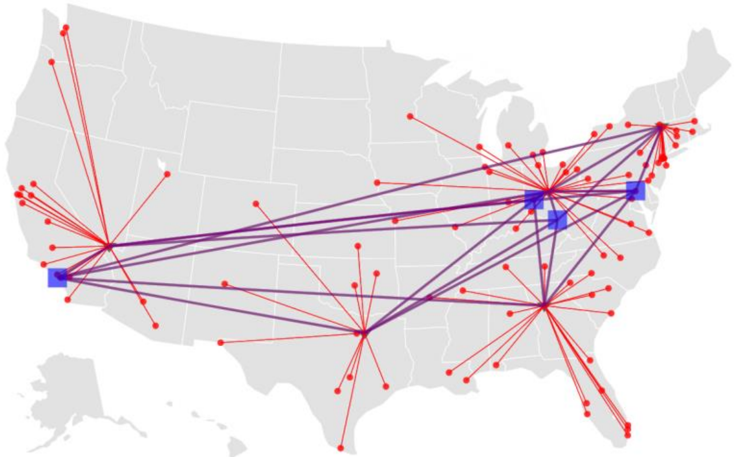
- The model elects to move the two warehouses (the new warehouse location in the east is both closer to customer points and closer to the other plants)
- The total transportation cost decreased to $196 million (from the original cost of $254 million):
  - The cost to get product from plants to warehouses went from $133 million to $94 million
  - The cost from warehouses to customers went from $121 million to $103 million
- The percentage of demand within 400 miles increased from 32% to 37%



Best 3 warehouse solutions



Best 4 warehouse solutions



Best 5 warehouse solutions

	Base case	Best 2	Best 3	Best 4	Best 5
Total cost (millions)	254	194	182	176	169
Plant to warehouse cost	133	94	102	107	109
Warehouse to customer	121	100	80	69	60
% service within 400 miles	32	37	63	71	72

- Simply moving the two warehouses to more optimal locations reduces transportation costs significantly. Adding more warehouses then only continues to reduce the cost.
- Regarding the warehouse fixed cost, going from the best three warehouses to the best four, a savings of $6 million is estimated. Therefore, if a new warehouse costs more than the $6 million, it would not make sense to open the fourth warehouse.
- The rate of savings decreases with each warehouse added to the network (the incremental value of additional warehouses decreases)
- The actual locations of the warehouses change within each solution. To convert the best three location solution to the best four location solution with the simple addition of one warehouse:
  - Pick the best three and add a best fourth warehouse later based on the original three locations as fixed within the model. This provides a good solution only for the best three, therefore it can be used if adding the fourth warehouse is not likely
  - Pick the best three out of the best four locations. It can be used if adding the fourth warehouse is somewhat likely
  - However it is good to test both cases and see whether the cost difference is significant.
- The savings of additional warehouses correlates directly with the savings in the warehouse-to-customer shipment costs
- The percentage of customers within 400 miles increases as the number of warehouses within the network increases

The validity of plant-to-warehouse rates depends on filling up the rail cars and trucks running on these lanes. If each plant is shipping to too many warehouses, however, it becomes more difficult to fill railcars and trucks. In this case, the ability to fill railcars and trucks is to be tested by analyzing the annual flow of product from each plant to each warehouse.

For instance:

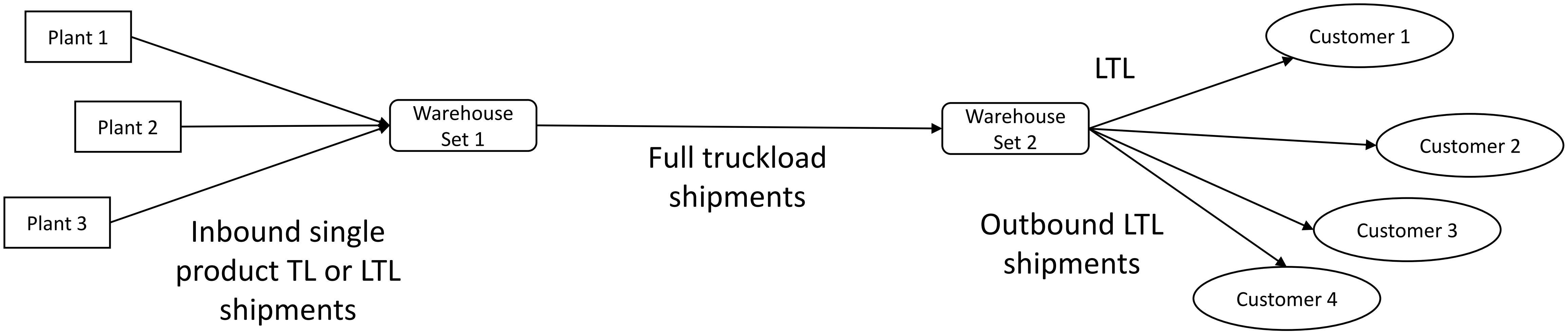
- 13,000,000 lbs of inbound product from plant to warehouses / 52 weeks = 250,000 lbs moving per week
- 250,000 lbs transferred to up to 5 warehouses = Average of 50,000 lbs to each warehouse
- 50,000 lbs > 45,000 lbs capacity of rail and truck transit modes
- railcars and trucks can be filled up with up to five total warehouse locations

To locate plants considering the source of raw materials:

- An approach similar to warehouse location studies is adopted
- The same trade-off between inbound and outbound transportation costs is investigated
- However, a plant can receive greater levels of inbound product compared to finished-good units being shipped out (e.g. two tons of raw material could be brought in for every ton of finished good that is shipped out)
- Therefore, the extra influence of total tons is to be considered in the analysis

The three-echelon modeling logic also applies to networks containing more than three echelons.
For a hub-and-spoke network:

- Distributors often specialize in selling products sourced from thousands of small vendors that are consolidated
- The most expensive parts of a distributor's supply chain are, most often, the collection of all the products from the numerous vendors and then making the final outbound delivery to the customer
- Therefore, multiple levels of warehouses are adopted:
    - the first set of warehouses collects products from the vendors
    - the second set is responsible for the final delivery to the customers
    - in the middle, the company can efficiently move product in full truckloads between warehouse sets



To locate warehouse set 1:

- A single or small number of central warehouses, when the vendors have enough volume to ship in full truckload to one or a few locations
- Many warehouses close to vendors and those vendors then ship short LTL shipments into the closest warehouse, when the vendors cannot fill a full truck even when shipping to just one location

To locate warehouse set 2:

- Numerous locations close to customers, since the company’s ability to transport product in full truckloads into these warehouses allows for the concentration on better service and shorter outbound LTL shipment distances to customers

For some companies, more than two echelons of warehouses are needed. For example, a major retailer that sells product and ship then directly to customers in a town:

- Consumer locations become people's houses, and orders may have multiple sizes:
  - For large items, companies will commonly use small trucks making multiple stops in a small geographic area
  - For the smaller items, these companies may use small-parcel delivery services
- In both cases, the objective is to keep the product on ground and in full trucks as long as possible to control costs. This requires local warehouses (third set). These locations are meant to receive product from regional warehouses and ease the high cost of small shipments deep into their markets.
- Because it can be expensive to operate a chain of fully stocked local warehouses, many firms, in practice, treat these warehouses as cross-docks:
  - Orders come in throughout the day for deliveries needed the next day.
  - After some cutoff time, the regional warehouse starts picking orders for the local warehouses. These orders are loaded onto a full truck.
  - Overnight, the truck drives to the local warehouse.
  - First thing in the morning, the product is moved from the regional warehouse truck onto the delivery vehicles.
- When firms ship small-parcel and are charged per package, a strategy called "zone-skipping" is adopted. The skipped "zones" are the zones set up by the small-parcel carriers. If there is enough volume, trucks can be filled up with the packages for a local market and driven straight to the parcel carrier's local sort facility. In this way, the product is given to the parcel carrier's delivery trucks.

The previous hub-and-spoke networks had fixed supply points.

More than three echelons are needed if optimal location of both the plants and the warehouses is simultaneously determined, leading to more complex problems.

In the hub-and-spoke example, the problem can be split into two problems:

- The first is locating the central warehouse close to the suppliers.
- Thanks to the ability to ship relatively low-cost full truckloads from the central to regional warehouses, the central warehouses can be located independent of customer locations.
- In parallel, the location of the regional warehouses does not depend on the central warehouses either. Instead, it depends solely on the location of the customers.

Capacity can be defined as the amount of output that a system is capable of achieving over a specific period of time.

Operations managers need to look at both resource inputs and product outputs. For planning purposes, real (or effective) capacity depends on what is to be produced.

While many industries measure and report their capacity in terms of outputs, those whose product mix is very uncertain often express capacity in terms of inputs.

Capacity must be stated relative to some period of time:

- Long range (greater than one year). Where productive resources (such as buildings, equipment, or facilities) take a long time to acquire or dispose of
- Intermediate range (monthly or quarterly plans for the next 6 to 18 months). Capacity may be varied by such alternatives as hiring, layoffs, new tools, minor equipment purchases, and subcontracting.
- Short range (less than one month). This is tied into the daily or weekly scheduling process and involves making adjustments to eliminate the variance between planned and actual output.

The objective of strategic capacity planning is to provide an approach for determining the overall capacity level of capital-intensive resources—facilities, equipment, and overall labor force size—that best supports the company’s long-term competitive strategy. The capacity level selected has a critical impact on the firm’s response rate, its cost structure, its inventory policies, and its management and staff support requirements.

The best operating level is the level of capacity for which the process was designed and thus is the volume of output at which average unit cost is minimized.
The capacity utilization rate reveals how close a firm is to its best operating level:

$$\frac{Capacity\ used}{Best\ operating\ level}$$

Potential exercise in the exam

The capacity utilization rate is expressed as a percentage.

Economies of scale: a plant gets larger and volume increases, the average cost per unit of output drops.
At some point, the size of a plant becomes too large and diseconomies of scale become a problem. For example:

- maintaining the demand required to keep the large facility busy may require significant discounting of the product
- using a few large-capacity pieces of equipment.

Capacity focus

A production facility works best when it focuses on a fairly limited set of production objectives.

A firm should select a limited set of tasks that contribute the most to corporate objectives. Typically, the focused factory would produce a specific product or related group of products. A focused factory allows capacity to be focused on producing those specific items.

Plant within a plant (PWP) is an area in a larger facility that is dedicated to a specific production objective (for example, product group). This can be used to operationalize the focused factory concept.

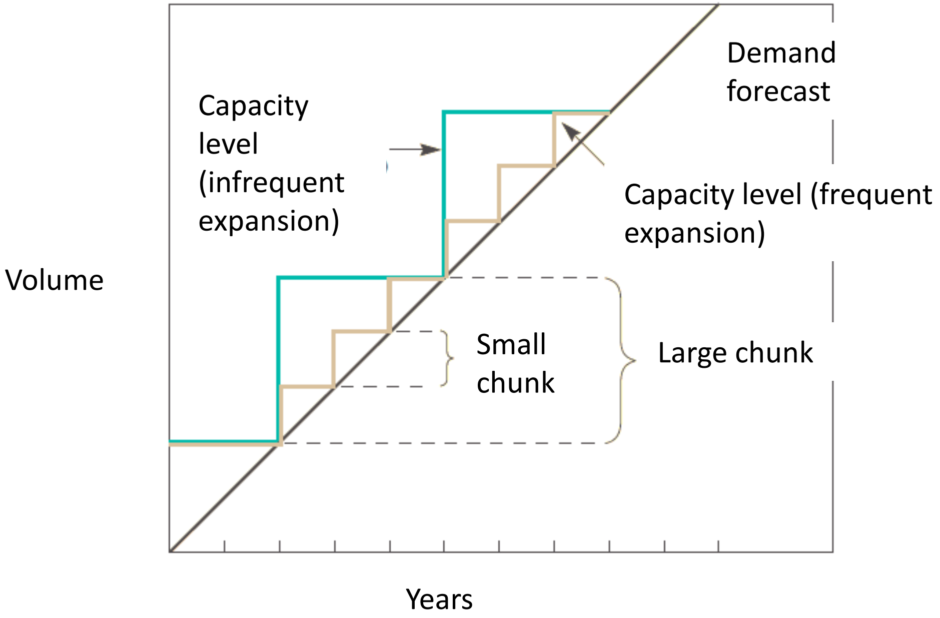
Capacity flexibility

The ability to rapidly increase or decrease production levels, or to shift production capacity quickly from one product or service to another. It can be achieved through:

- Flexible plants
- Flexible processes, gathered by flexible manufacturing systems on the one hand and simple, easily set up equipment on the other
- Flexible Workers
- Strategies that use the capacity of other organizations

Issues to be considered when adding or decreasing capacity:

- Maintaining System Balance. In a perfectly balanced plant the output of one stage is the input of the next one. However, the best operating levels for each stage generally differ and there could be variability in product demand. To deal with imbalance:
  - Add capacity to stages that are bottlenecks
  - Include buffer inventories in front of the bottleneck stage to ensure it always has something to work on
  - Increase the facilities of one department on which another is dependent
- Frequency of Capacity Additions. Both upgrading too frequently and upgrading too infrequently are expensive
- External Sources of Operations and Supply Capacity. Two common strategies used by organizations are outsourcing and sharing capacity.
- Decreasing Capacity. Shedding capacity in response to decreased demand can create significant problems for a firm.



To estimate capacity requirements, the demands for individual product lines, individual plant capabilities, and allocation of production throughout the plant network must be addressed according to the following steps:

- Use forecasting techniques to predict sales for individual products within each product line.
- Calculate equipment and labor requirements to meet product line forecasts.
- Project labor and equipment availabilities over the planning horizon.

A capacity cushion is an amount of capacity in excess of expected demand.

When a firm's design capacity is less than the capacity required to meet its demand, it is said to have a negative capacity cushion.

Example

A produces two brands of salad dressings: Paul's and Newman's. Each is available in bottles and single-serving plastic bags. Management would like to determine yearly equipment and labor requirements for its packing operation for the next five years.

	Year 1	Year 2	Year 3	Year 4	Year 5
Paul's					
Bottles	60000	100000	150000	200000	250000
Plastic bags	1000	200000	300000	400000	500000
Newman's					
Bottles	75000	85000	95000	97000	98000
Plastic bags	200000	400000	600000	650000	680000

Potential exercise in the exam

The company has:

- three machines that can each package 150,000 bottles each year (each machine has two operators)
- five machines that can each package 250,000 plastic bags per year (each of these machines has three operators)

Will the company have enough yearly packaging capacity to meet future demand?

- Step 1. Use forecasting techniques to predict sales for individual products within each product line (in the previous table).
- Step 2. Calculate equipment and labor requirements to meet product line forecasts. Currently, 15 plastic bag machine operators are available. Total product line forecasts can be calculated from the preceding table by adding the yearly demand for bottles and plastic bags.

	Year 1	Year 2	Year 3	Year 4	Year 5
Bottles	135000	185000	245000	297000	348000
Plastic bags	300000	600000	900000	1050000	1180000

Potential exercise in the exam

Calculate equipment and labor requirements for the current year:

- The total available capacity for packaging bottles is 450,000/year (3 machines × 150,000 each)
- 135,000/450,000 = 0.3 or 30 percent of the available capacity is used for the current year, or 135,000/150,000 = 0.9 machine
- 300,000/1,250,000 = 0.24 or 24 percent of the available capacity for plastic bags is needed for the current year, or 300,000/250,000 = 1.2 machines.
- The labor requirement for year 1 is: 0.9 bottle machine × 2 operators = 1.8 operators and 1.2 bag machines × 3 operators = 3.6 operators

	1	2	3	4	5
Bottle Operation Percentage machine capacity utilized	30%	41%	54.4%	66%	77.3%
Machine requirement	0.9	1.23	1.63	1.98	2.32
Labor requirement	1.8	2.46	3.26	3.96	4.64
Plastic Bag Operation Percentage machine capacity utilized	24%	48%	72%	84%	94%
Machine requirement	1.2	2.4	3.6	4.2	4.7
Labor requirement	3.6	7.2	10.8	12.6	14.1

Potential exercise in the exam

- Step 3. Project labor and equipment availabilities over the planning horizon (5 years)

A positive capacity cushion exists for all five years relative to the availability of machines because the available capacity for both operations always exceeds the expected demand.

A decision tree is a schematic model of the sequence of steps in a problem and the conditions and consequences of each step:

- The tree format helps not only in understanding the problem but also in finding a solution.
- Decision trees are composed of decision nodes with branches extending to and from them. Squares represent decision points and circles represent chance events. Branches from decision points show the choices available to the decision maker; branches from chance events show the probabilities for their occurrence.
- To solve decision tree problems:
  - work from the end of the tree backward to the start of the tree
  - calculate the expected values at each step.
  - once the calculations are made, prune the tree by eliminating from each decision point all branches except the one with the highest payoff.
  - this process continues to the first decision point.

Example

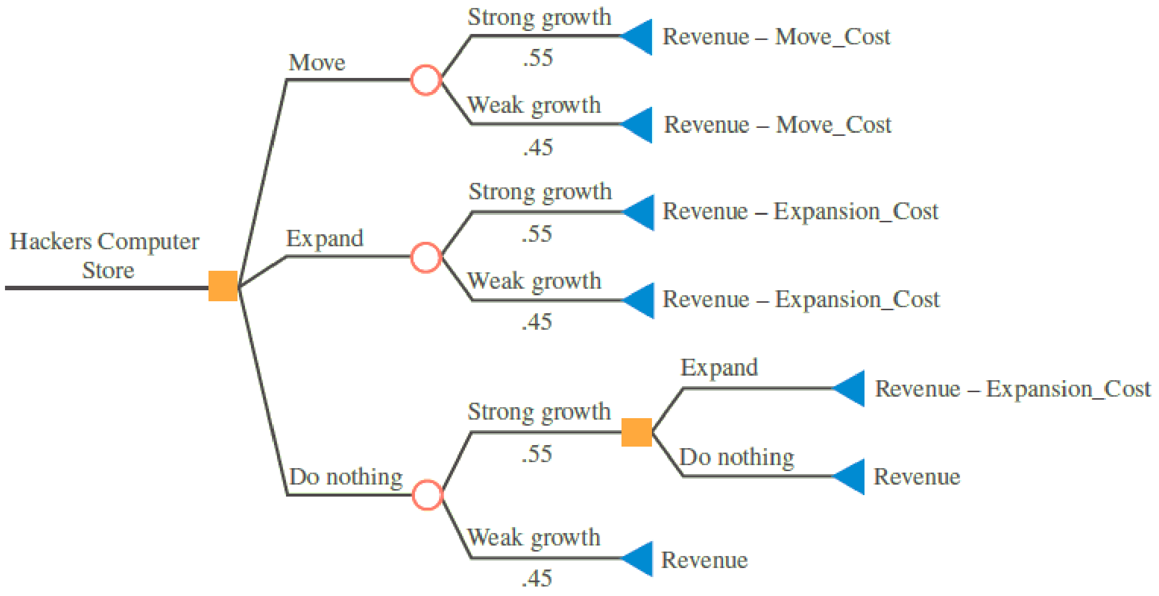
The firm owning a computer store must decide one of the following intervention for the next five years:

- enlarge his current store
- locate at a new site
- wait and do nothing.

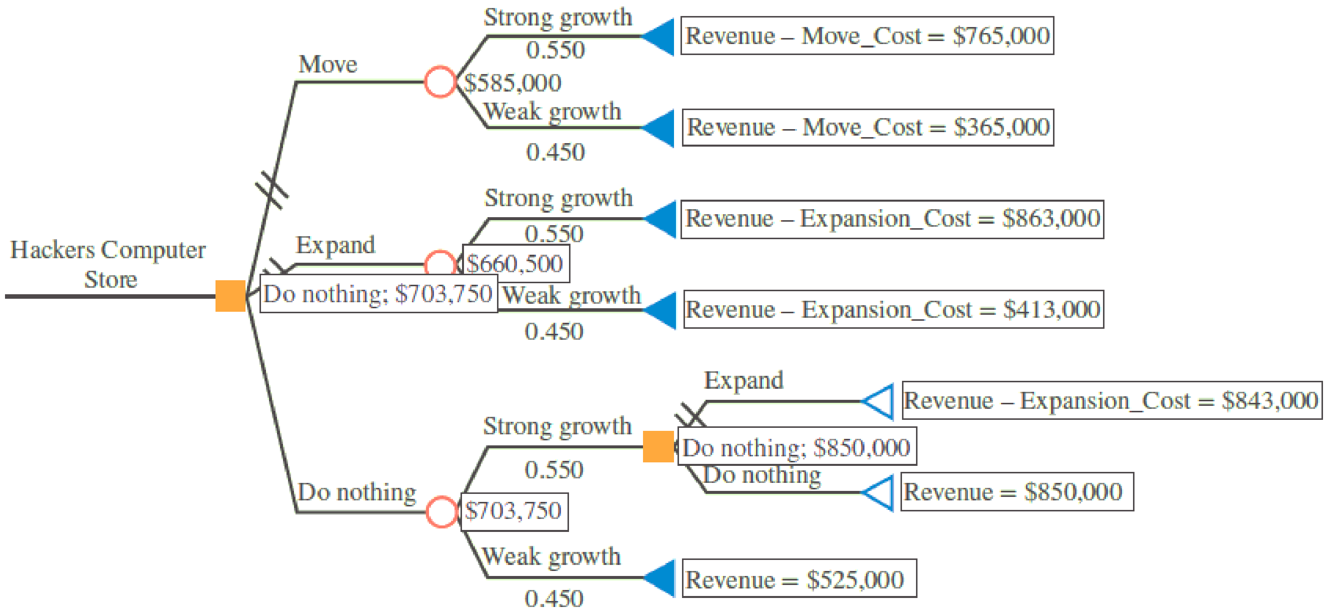
The assumptions and conditions are as follows:

- Strong growth as a result of the increased population of computer fanatics from the new electronics firm has a 55 percent probability.
- Strong growth with a new site would give annual returns of $195,000 per year. Weak growth with a new site would mean annual returns of $115,000.
- Strong growth with an expansion would give annual returns of $190,000 per year. Weak growth with an expansion would mean annual returns of $100,000.
- At the existing store with no changes, there would be returns of $170,000 per year if there is strong growth and $105,000 per year if growth is weak.
- Expansion at the current site would cost $87,000.
- The move to the new site would cost $210,000.
- If growth is strong and the existing site is enlarged during the second year, the cost would still be $87,000.
- Operating costs for all options are equal.

A decision tree is built. There are two decision points (shown with the square nodes) and three chance occurrences (round nodes).



Alternative	Revenue	Cost	Value
Move to new location, strong growth	\$195,000 × 5 yrs	\$210,000	\$765,000
Move to new location, weak growth	\$115,000 × 5 yrs	\$210,000	\$365,000
Expand store, strong growth	\$190,000 × 5 yrs	\$87,000	\$863,000
Expand store, weak growth	\$100,000 × 5 yrs	\$87,000	\$413,000
Do nothing now, strong growth, expand next year	\$170,000 × 1 yr + \$190,000 × 4 yrs	\$87,000	\$843,000
Do nothing now, strong growth, do not expand next year	\$170,000 × 5 yrs	\$0	\$850,000
Do nothing now, weak growth	\$105,000 × 5 yrs	\$0	\$525,000



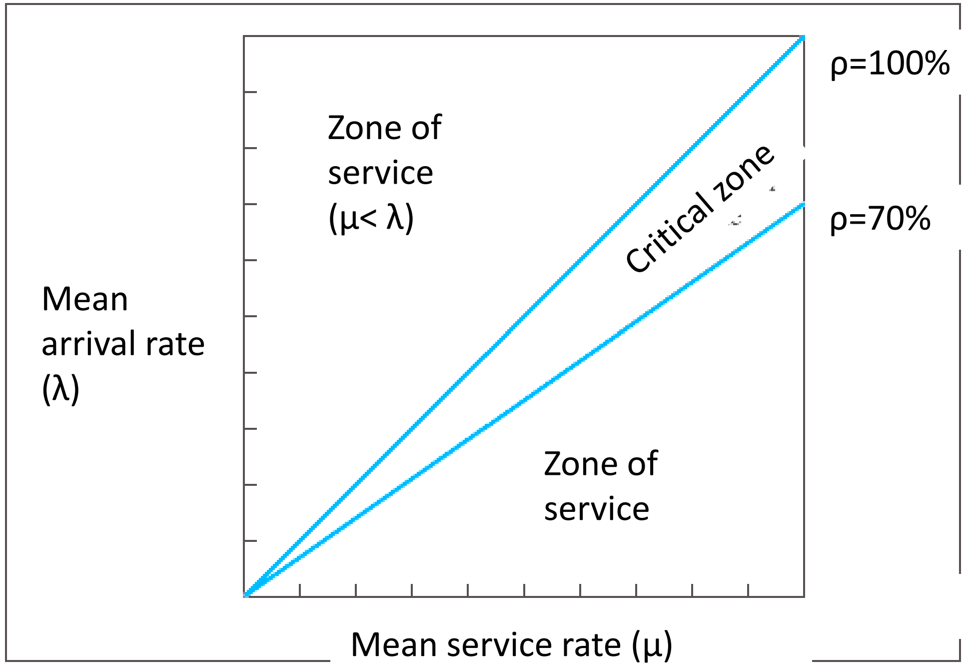
Working from the rightmost alternatives, which are associated with the decision of whether to expand, the alternative of doing nothing has a higher value than the expansion alternative. We therefore eliminate the expansion in the second year alternatives.

The expected values associated with current decision alternatives are calculated by multiplying the value of the alternative by its probability and sum the values.

Service capacity planning is subject to many of the same issues as manufacturing capacity planning, and facility sizing can be done in much the same way, however service capacity is more time and location-dependent, it is subject to more volatile demand fluctuations, and utilization directly impacts service quality:

- Unlike goods, services cannot be stored for later use. As such, in services, managers must consider time as one of their supplies. The capacity must be available to produce a service when it is needed.
- The capacity to deliver the service must first be distributed to the customer (either physically or through some communications medium), then the service can be produced.
- The volatility of demand on a service delivery system is much higher than that on a manufacturing production system since:
  - services cannot be stored (inventory cannot smooth the demand as in manufacturing).
  - the customers interact directly with the production system and have different needs, levels of experience with the process, and may require a different number of transactions. This contributes to greater variability in the processing time required for each customer and hence greater variability in the minimum capacity needed.
  - it is directly affected by consumer behavior.

Planning capacity levels for services must consider the day-to-day relationship between service utilization and service quality.

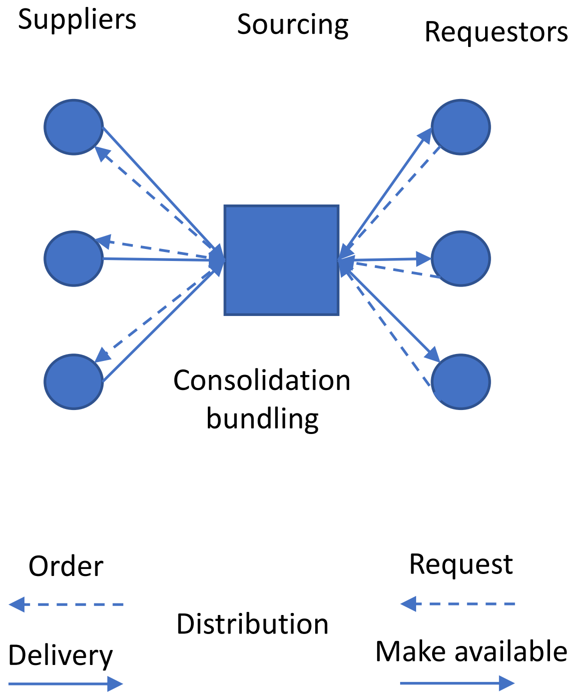


- The arrival rate is the average number of customers that come to a facility during a specific period of time.
- The service rate is the average number of customers that can be processed over the same period of time when the facility is operating at maximum capacity.
- The best operating point is near 70 percent of the maximum capacity (enough to keep servers busy but allows enough time to serve customers individually and keep enough capacity in reserve so as not to create too many managerial headaches).
- In the critical zone, customers are processed through the system, but service quality declines.
- Above the critical zone, where customers arrive at a rate faster than they can be served, the line builds up and it is likely that many customers may never be served.
- The optimal utilization rate is very context specific. Low rates are appropriate when both the degree of uncertainty and the stakes are high. Relatively predictable services can plan to operate much nearer to 100 percent utilization.

Purchasing is the process of procuring the proper requirement, at the necessary time, for the lowest possible cost from a reliable source.
Procurement covers both acquisitions from third parties and from in-house providers.
Sourcing is the development and management of supplier relationships to acquire goods and services in a way that aids in achieving the immediate needs of the business.

The sourcing process connects the supply side with the demand side. The content and objectives of sourcing activities are:

- Determination of material requirements
- Order management
- Supplier base research, observation, and analysis
- Make or buy decisions
- Supplier management



Make-or-buy

Make-or-buy is a strategic decision that determines if the sourcing objects are internally made or externally sourced. Such strategic decisions might also be related to the question of core competencies.

It is closely linked to decisions around production strategy. The sourcing team will need to decide jointly with manufacturing experts whether external purchasing ("buy") might be the economically preferable solution or if the respective part should be produced in-house ("make").

Outsourcing

Outsourcing is the act of moving some of a firm's internal activities and decision responsibility to outside providers. Unlike purchasing and consulting contracts, not only the activities are transferred, but also resources that make the activities occur, including people, facilities, equipment, technology, other assets, and the responsibilities for making decisions over certain elements of the activities are transferred.

Sourcing strategies can be classified according to three basic features:

- number of suppliers (single, dual, and multiple sourcing). The task is to manage the supplier base with the objective of determining the right number of suppliers. By reducing the supplier base, larger volumes can be ordered from one supplier (single sourcing strategy) with the objective of generating volume bundling (scale) effects. However, a multiple sourcing strategy which both better balance the global flows of material and provide risk reduction.
- geographical supplier distribution (local, national, international, and global sourcing). There are two extremes, starting from local sourcing (limited number of suppliers, but same norms, language, currency, shortest distance, and thus fastest reaction time in case of changes) versus global sourcing (offering the broadest supplier base, but suffering from e.g., long distances, different currencies, norms standards). Between these two extremes, there are also opportunities for following a national or continental sourcing strategy
- sourcing principles (sourcing on-stock, Just-In-Time, particular sourcing).

	Local sourcing	Global sourcing
Advantages	• Same norms/standards • Easy to reach/short distances • Same culture, same currency, samepolitical climate • Good basis for JIT deliveries • Lower disruption risks for overallSC	• Broadest variety of available vendors • Largest portfolio of products orservices • Best opportunities to compare andnegotiate with suppliers due to broadest supplier base
Disadvantages	• (Very) limited supplier base or nosupplier base • Possibly limited bargaining powerof buyer because of limitations on supplier side	• Longer travel and transportationtimes • Longer response times in the event ofchanges • Possibly larger lot sizes • Potentially different norms/standards • Different cultures, currencies, andpolitical uncertainty • Higher disruption risks for overall SC

Total Cost of Ownership

It is an estimate of the cost of an item that includes all the costs related to its procurement and use, including any related costs in disposing of the item after it is no longer useful.

The costs can be categorized into three broad areas:

- acquisition costs. They are the initial costs associated with the purchase of materials, products, and services.
- ownership costs. They are incurred after the initial purchase and are associated with the ongoing use of the product or material.
- post-ownership costs. They include salvage value and disposal costs.

These costs can be estimated as cash inflows (the sale of used equipment, etc.) or outflows (such as purchase prices, demolition of an obsolete facility, etc.).

Example: purchase of a copy machine

- initial cost: $120,000
- Income: $40,000 per year.
- Supplies: $7,000 per year
- It needs to be overhauled during year 3 at a cost of $9,000.
- It has a salvage value of $7,500 when it is sold at year 6.

Laying these costs out over time can lead to the use of net present value analysis to evaluate the decision (Discount factor of 20%).

Year	Now	1	2	3	4	5	6
Initial Investment	–$120,000						
Manufacturer required overhaul				–$9,000			
Cash inflows from using the machine		$ 40,000	$ 40,000	$ 40,000	$ 40,000	$ 40,000	$ 40,000
Supplies needed to use the machine		–$7,000	–$7,000	–$7,000	–$7,000	–$7,000	–$7,000
Salvage value							$ 7,500
Total of annual streams	–$120,000	$ 33,000	$ 33,000	$ 24,000	$ 33,000	$ 33,000	$ 40,500
Discount factor from Appendix C $1/(1+.2)^{Year}$	1.000	0.833	0.694	0.579	0.482	0.402	0.335
Present value – yearly	–$120,000	$ 27,500	$ 22,917	$ 13,889	$ 15,914	$ 13,262	$ 13,563
Present value	–$12,955						

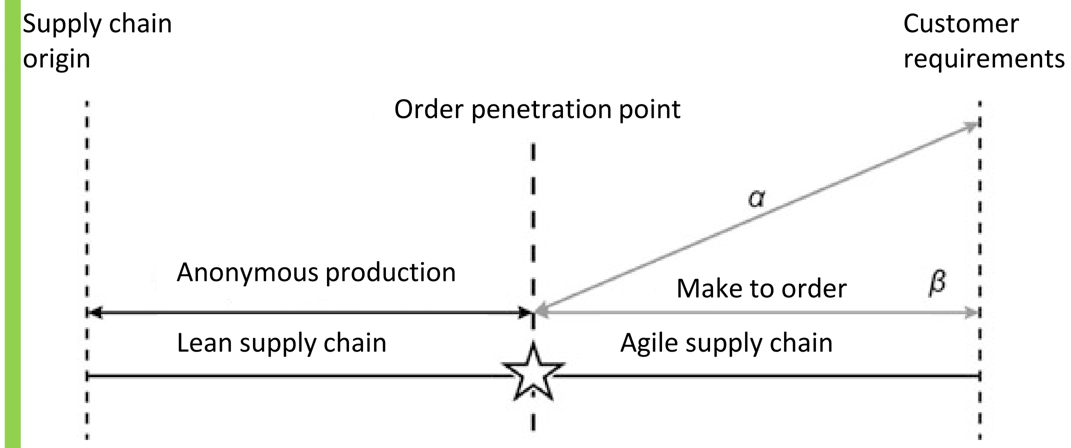
The aim is to determine the right production strategy for the supply chain in order to combine the advantages of mass production and customization. In order to improve responsiveness without holding huge finished goods inventories, postponement and modularization can be used:

- Postponement is an organizational concept whereby some of the activities in the supply chain are not performed until customer orders are received. Differentiation of a generic product into a specific end-product is shifted closer to the consumer by postponing identity changes, such as assembly or packaging, until the last possible supply chain location. This allows safety stock to be held as one generic product instead of multiple specific end-products. Postponement leads to a lower safety stock requirement and a lower risk of obsolescence of end-products.
- Modularization is the building a complex product or process from smaller subsystems that can be designed independently yet function together as a whole.

Following the push/pull view, the supply chain processes can be divided into two categories based on whether they are executed in response to a customer order (downstream supply chain) or in anticipation of customer orders (make-to-stock, upstream supply chain):

- Pull processes are initiated in response to a customer order. The pull advantages are responsiveness, and a high degree of customer-oriented product individualization.
- Push processes are initiated and performed in anticipation of customer orders. The advantages of this are economy of scale (low manufacturing and transportation costs), flexibility (high level of inventory), and short supply times.

The main idea behind postponement is to delay product differentiation at a point closer to the customer called order penetration point (OPP).



In the case of a customer inquiry or order, the goods can be delivered from a general inventory holding unit (trajectory α) or customized according to the inquiry or order (trajectory β). Upstream of the OPP, the processes are designed to be lean. Downstream the processes are designed to be agile.

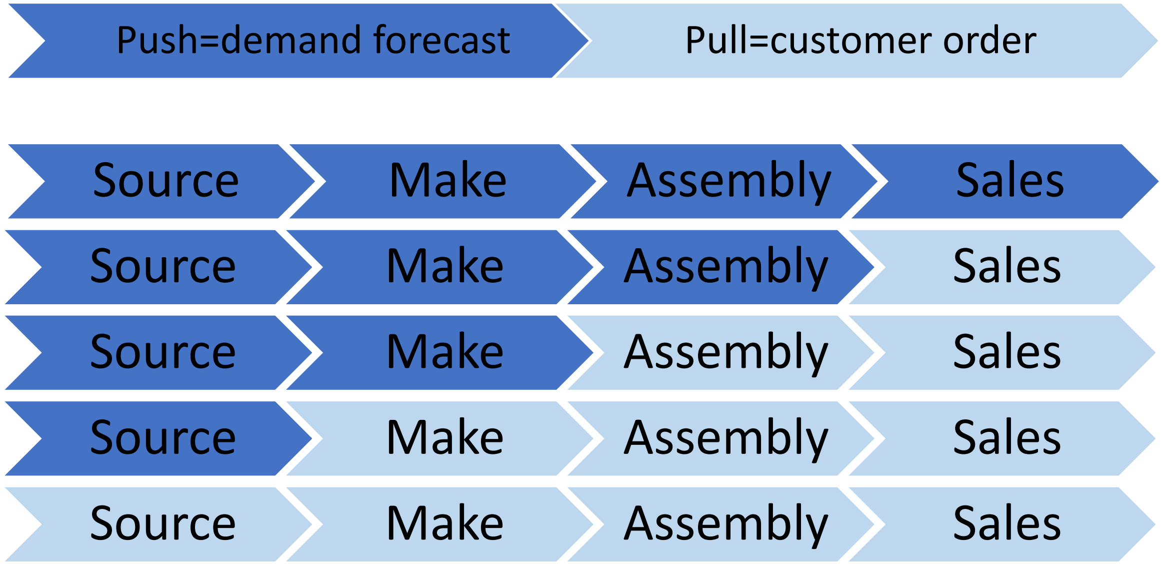
Three categories are primary relevant to the OPP location:

- Delay of product differentiation (postponement)
- Design and developing standard or generic configurable products that can be customized quickly (product modularization)
- Implementation of specific collaboration strategies, inventory strategies, buffers, etc. to fulfil service level objectives (agility).

In practice, a selected postponement strategy determines the OPP location. The further upstream the OPP is located in the SC, the more manufacturing and logistics activities are postponed. In general, the aim is to postpone manufacturing and logistics as much as possible and maintain high customer service standards.

- With pull processes, execution is initiated in response to a customer order. Therefore, at the time of execution of a pull process, customer demand is known with certainty,
- With push processes, execution is initiated in anticipation of customer orders. Therefore, at the time of execution of a push process, demand is not known and must be forecasted.

The push/pull boundary is the place of the OPP that separates push processes from pull processes. A product is kept as long as possible in a generic state.



According to the placement of the OPP, the following production strategies can be determined:

- Make-to-stock (MTS)
- Distribute-to-order (DTO)/Configure-to-order (CTO)
- Assemble-to-order (ATO)
- Make-to-order (MTO)
- Engineer-to-order (ETO).

- MTS strategy is typical for mass production of standard products (sugar, socks). MTS advantages are low unit production costs and higher capacity utilization because of scale effect and quantity flexibility regarding demand fluctuations resulting from product standardization. Disadvantages of MTS are higher inventory and lower production flexibility.
- DTO or CTO strategies allow a small degree of customer individualization. It might have customer-specific packaging (as in the pharmaceutical industry) or some individual items added to a standard manufactured product. Advantages of DTO are higher flexibility regarding product structure and lower inventory.
- ATO presumes a higher degree of product individualization (directly within the manufacturing). Assembly is performed individually for each customer from a set of standard modules. Advantages and shortcomings of the ATO are similar to the DTO/CTO, but with higher investments in the process, product flexibility, and lower inventory.
- MTO strategy presumes also a possibility of customer individualization at the module/component level (also that the modules and components can be manufactured individually for each particular customer). Advantages and shortcomings of MTO are similar to ATO, but with even higher investments in flexibility and lower inventory.
- ETO has the highest degree of customer individualization. ETO is reasonable for complex products of high value where customer individualization is crucial.

The Lost-Sales Analysis is used to investigate the issue of how to determine what the right production strategy and OPP location is in the supply chain.

Two strategies are considered:

- MTS and delivering from a general inventory holding unit and
- introducing the agile MTO part downstream from the general inventory holding unit.

The OPP location can be determined by a comparison of the financial results of the two strategies (with and without OPP). ,
A quantitative estimation of the OPP location can be found by relating the above-mentioned strategies to each other:

$$D = \frac{R_a k_a - (C_u^a + C_d^a + P^a + L^a)}{R - (C + kP + kL)} \rightarrow \max[t_{0;T}]$$

Different OPP locations can be investigated and the best one with the maximum value D > 1 is selected on the basis of the formula.

D is the index to characterize the efficiency of the OPP location,
R is the revenue in case of make-to-stock,
R_a is the revenue when agility is introduced downstream from the general inventory holding unit,
C are the supply chain costs in case of make-to-stock,
C_u^a are the supply chain costs upstream from the OPP when agility is introduced, C_d^a are the SC costs downstream from the OPP when agility is introduced, P are penalties of not-fulfilled contracts in case of make-to-stock,
P_a are penalties of not-fulfilled contracts when agility is introduced,
L are losses of rejected customer's inquiries in case of make-to-stock,
La are losses of rejected customer's inquiries when agility is introduced,
k, ka are correcting coefficients to take into account future increases in sales due to an increase in responsiveness when agility is introduced,
t is instants of time within the whole SC cycle [t_0; T].

Purpose of the forecast:

- Strategic forecasts Medium and long-term forecasts that are used for decisions related to strategy and aggregate demand.
- Tactical forecasts Short-term forecasts used for making day-to-day decisions related to meeting demand.

A perfect forecast is virtually impossible since too many factors in the business environment cannot be predicted with certainty. Therefore, rather than search for the perfect forecast, it is far more important to establish the practice of continual review of forecasts and to learn to live with inaccurate forecasts.

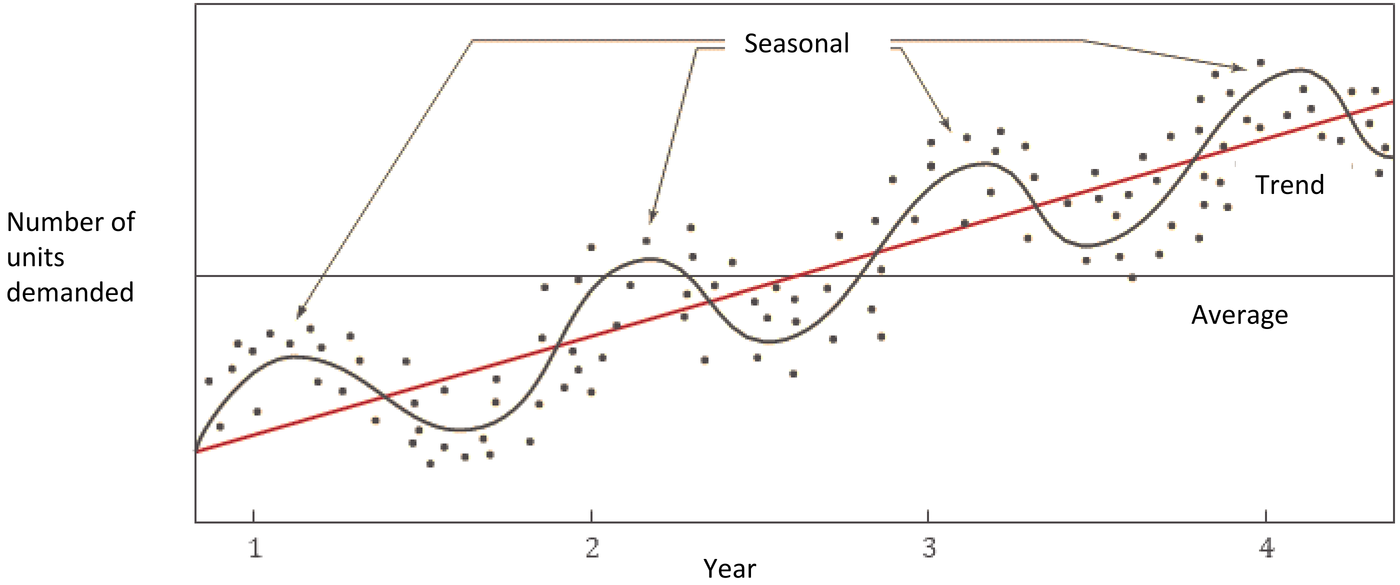
When forecasting, a good strategy is to use two or three methods and look at them from a common-sense view.

Forecasting methods:

- Qualitative. They are primarily subjective and rely on human judgment. They are most appropriate when little historical data are available or when experts have market intelligence that may affect the forecast.
- Time series. They use historical demand to make a forecast. They are based on the assumption that past demand history is a good indicator of future demand.
- Causal. They assume that the demand forecast is highly correlated with certain factors in the environment (the state of the economy, interest rates, etc.).
- Simulation. They methods imitate the consumer choices that give rise to demand to arrive at a forecast.

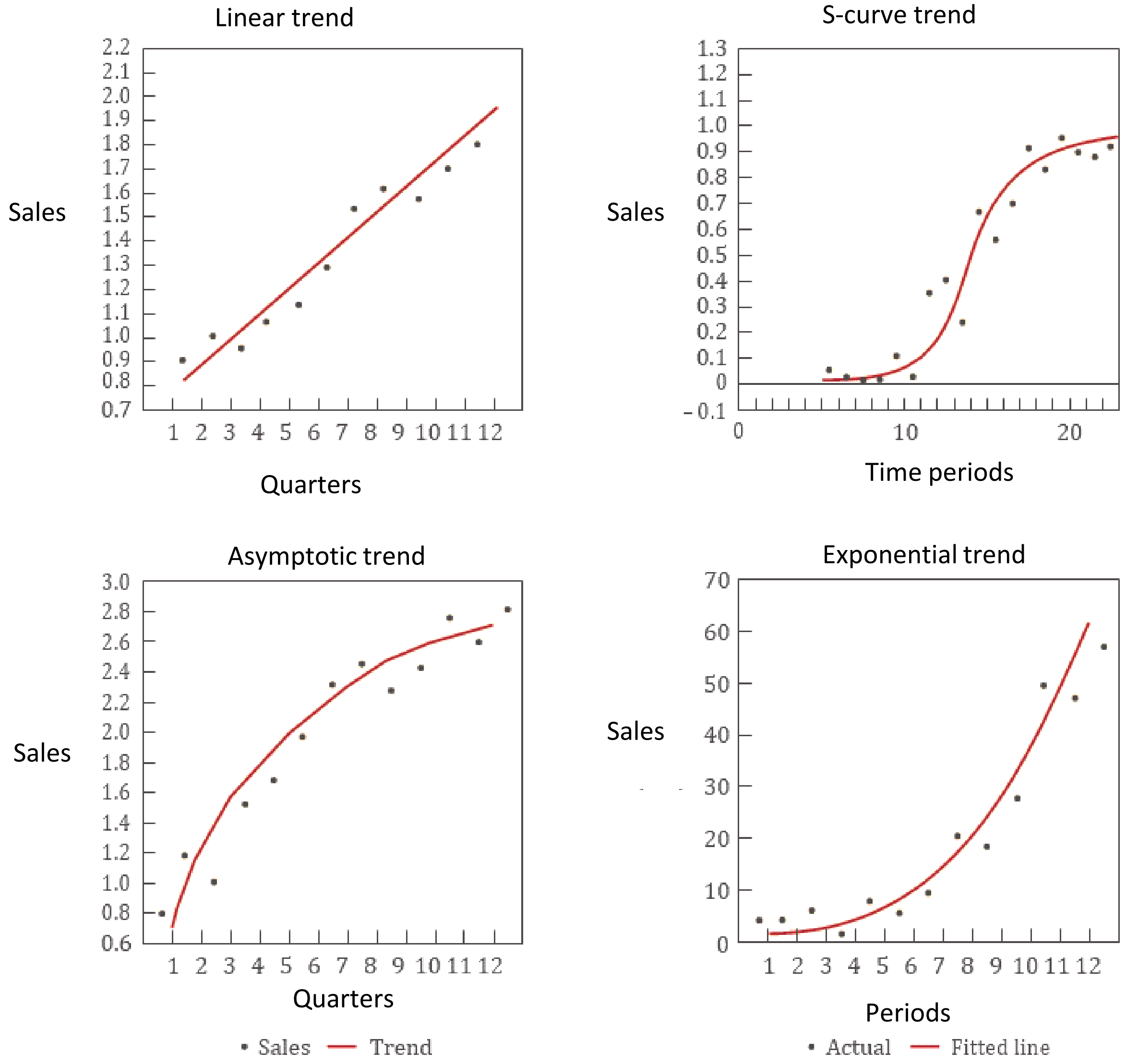
Demand components:

- Systematic part:
  - average demand for the period
  - cyclical elements. Cyclical factors are more difficult to determine because the time span may be unknown or the cause of the cycle may not be considered. Cyclical influence on demand may come from such occurrences as political elections, war, economic conditions, or sociological pressures.
  - autocorrelation. It denotes the persistence of occurrence (the value expected at any point is highly correlated with its own past values)



- seasonal element. The predictable seasonal fluctuations in demand
- a trend. It is the rate of growth or decline in demand for the next period. A widely used forecasting method plots data and then searches for the curve pattern (such as linear, S-curve, asymptotic, or exponential) that fits best

- random variation. It is the unexplained portion of demand and it is assumed to be purely random chance. All a company can predict is the random component's size and variability, which provides a measure of forecast error that measures the difference between the forecast and actual demand. On average, a good forecasting method has an error whose size is comparable to the random component of demand.



Steps for demand forecasting:

- Understand the objective of forecasting (e.g. how much of a particular product to make, how much to inventory, and how much to order)
- Integrate demand planning and forecasting throughout the supply chain.
- Identify the major factors that influence the demand forecast:
  - On the demand side, a company must ascertain whether demand is growing or declining or has a seasonal pattern.
  - On the supply side, a company must consider the available supply sources to decide on the accuracy of the forecast desired.
  - On the product side, a firm must know the number of variants of a product being sold and whether these variants substitute for or complement one another.
- Forecast at the appropriate level of aggregation. It is important to forecast at a level of aggregation that is appropriate, given the supply chain decision that is driven by the forecast.
- Establish performance and error measures for the forecast. These measures should be linked to the objectives of the business decisions based on these forecasts.

Time Series Analysis

Time series forecasting models try to predict the future based on past data.

Time horizon is relative to the context in which they are used:

- short term (under three months); for tactical decisions such as replenishing inventory or scheduling employees in the near term. They are especially good for measuring the current variability in demand
- medium term (three months to two years); for planning a strategy for meeting demand over the next six months to a year and a half. They are useful for capturing seasonal effects
- long term (greater than two years). They detect general trends and are especially useful in identifying major turning points

Forecasting Method	Amount of Historical Data	Data Pattern	Forecast Horizon
Simple moving average	6 to 12 months; weekly data are often used	Stationary only (i.e., no trend orseasonality)	Short
Weighted moving average andsimple exponential smoothing	5 to 10 observations needed to start	Stationary only	Short
Exponential smoothing with trend	5 to 10 observations needed to start	Stationary and trend	Short
Linear regression	10 to 20 observations	Stationary, trend, and seasonality	Short to medium
Trend and seasonal models	2 to 3 observations per season	Stationary, trend, and seasonality	Short to medium

The selection of forecasting model depends on:

- Time horizon to forecast
- Data availability
- Accuracy required
- Size of forecasting budget
- Availability of qualified personnel

Simple Moving Average

When demand for a product is neither growing nor declining rapidly, and if it does not have seasonal characteristics, a moving average can be useful in removing the random fluctuations for forecasting. The idea here is to simply calculate the average demand over the most recent periods. Each time a new forecast is made, the oldest period is discarded in the average and the newest period included.

Potential exercise in the exam

Week	Demand	3 Week	9 Week
1	800		
2	1,400		
3	1,000		
4	1,500	1,067	
5	1,500	1,300	
6	1,300	1,333	
7	1,800	1,433	
8	1,700	1,533	
9	1,300	1,600	
10	1,700	1,600	1,367
11	1,700	1,567	1,467
12	1,500	1,567	1,500
13	2,300	1,633	1,556
14	2,300	1,833	1,644
15	2,000	2,033	1,733

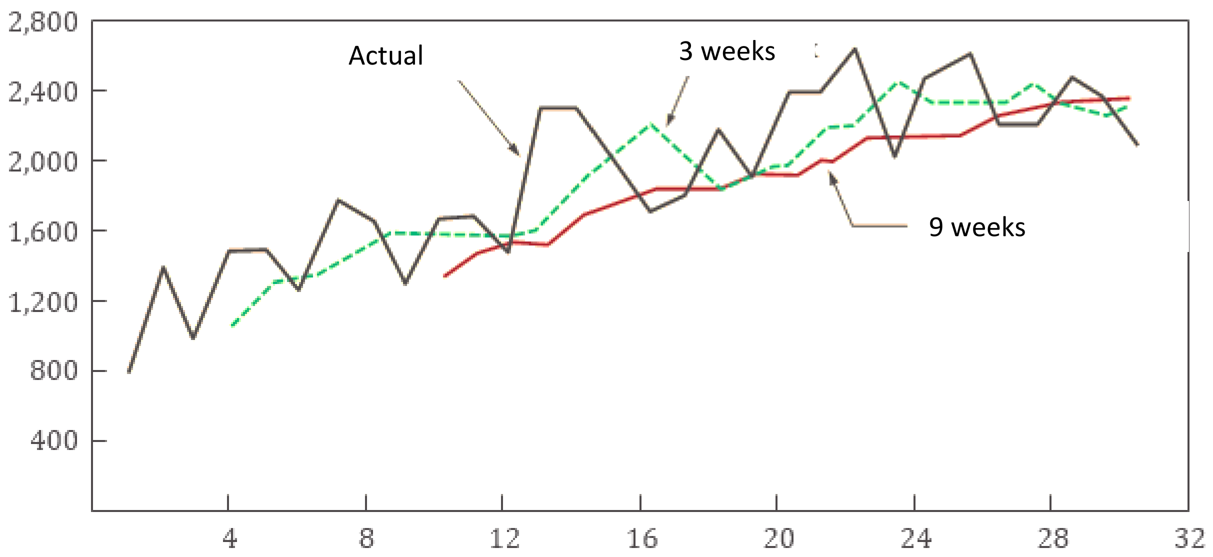
Week	Demand	3 Week	9 Week
16	1,700	2,200	1,811
17	1,800	2,000	1,800
18	2,200	1,833	1,811
19	1,900	1,900	1,911
20	2,400	1,967	1,933
21	2,400	2,167	2,011
22	2,600	2,233	2,111
23	2,000	2,467	2,144
24	2,500	2,333	2,111
25	2,600	2,367	2,167
26	2,200	2,367	2,267
27	2,200	2,433	2,311
28	2,500	2,333	2,311
29	2,400	2,300	2,378
30	2,100	2,367	2,378

- Selecting the period length should be dependent on how the forecast is going to be used (in the case of a medium-term forecast of demand for planning a budget, monthly time periods might be more appropriate, whereas if the forecast were being used for a short- term decision related to replenishing inventory, a weekly forecast might be more appropriate).
- The number of periods to use in the forecast can also have a major impact on the accuracy of the forecast. As the moving average period becomes shorter, and fewer periods are used, and there is more oscillation, there is a closer following of the trend. Conversely, a longer time span gives a smoother response, but lags the trend.
- The formula is:

$$F_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \ldots + A_{t-n}}{n}$$

Potential exercise in the exam

F_t = Forecast for the coming period
n = Number of periods to be averaged
A_{t-1} = Actual occurrence in the past period
A_{t-2}, A_{t-3}, and A_{t-n} = Actual occurrences two periods ago, three periods ago, and so on, up to n periods ago



- The growth trend levels off at about the 23rd week. The three- week moving average responds better in following this change than the nine-week average, although overall, the nine-week average is smoother.

The main disadvantage in calculating a moving average is that all individual elements must be carried as data because a new forecast period involves adding new data and dropping the earliest data.

Weighted Moving Average

A forecast made with past data where more recent data is given more significance than older data. Whereas the simple moving average assigns equal importance to each component of the moving average database, a weighted moving average allows any weights to be placed on each element, provided, of course, that the sum of all weights equals 1.

Example: a department store may find that, in a four-month period, the best forecast is derived by using 40 percent of the actual sales for the most recent month, 30 percent of two months ago, 20 percent of three months ago, and 10 percent of four months ago.

Potential exercise in the exam

Month 1	Month 2	Month 3	Month 4	Month 5
100	90	105	95	?

The forecast for month 5 would be

$$F_5 = 0.40(95) + 0.30(105) + 0.20(90) + 0.10(100) = 38 + 31.5 + 18 + 10 = 97.5$$

The formula for a weighted moving average is:

$$F_t = w_1 A_{t-1} + w_2 A_{t-2} + \cdots + w_n A_{t-n}$$

w_1 = Weight to be given to the actual occurrence for the period $t - 1$
w_2 = Weight to be given to the actual occurrence for the period $t - 2$
w_n = Weight to be given to the actual occurrence for the period $t - n$
n = Total number of prior periods in the forecast

The sum of all the weights must equal 1:

$$\sum_{i=1}^{n} w_i = 1$$

Potential exercise in the exam

Example: Suppose sales for month 5 actually turned out to be 110. Then, the forecast for month 6 would be:

$$F_6 = 0.40(110) + 0.30(95) + 0.20(105) + 0.10(90) = 44 + 28.5 + 21 + 9 = 102.5$$

Experience and trial and error are the simplest ways to choose weights. As a general rule, the most recent past is the most important indicator of what to expect in the future, and, therefore, it should get higher weighting.

Exponential Smoothing

A time series forecasting technique using weights that decrease exponentially $(1 - \alpha)$ for each past period.
The idea is that the importance of data diminishes as the past becomes more distant.
Exponential smoothing is the most used of all forecasting techniques:

- Exponential models are surprisingly accurate.
- Formulating an exponential model is relatively easy.
- The user can understand how the model works.
- Little computation is required to use the model.
- Computer storage requirements are small because of the limited use of historical data.
- Tests for accuracy as to how well the model is performing are easy to compute.

Potential exercise in the exam

Three pieces of data are needed to forecast the future:

- the most recent forecast
- the actual demand that occurred for that forecast period
- a smoothing constant alpha (α). It controls the speed of reaction to differences between forecasts and actual demand. The value for the constant is determined both by the nature of the product and by the manager's sense of what constitutes a good response rate:
  - if a firm produced a standard item with relatively stable demand, the reaction rate to differences between actual and forecast demand would tend to be small (5 or 10 percentage points).
  - if the firm were experiencing growth, it would be desirable to have a higher reaction rate (15 to 30 percentage points), to give greater importance to recent growth experience.
  - The more rapid the growth, the higher the reaction rate should be.
  - To keep the forecasts about the same as the simple moving average, α is approximated by $2 \div (n + 1)$, where n is the number of time periods in the corresponding simple moving average.

Potential exercise in the exam

The equation for a single exponential smoothing forecast is:

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$

F_t = The exponentially smoothed forecast for period t
F_{t-1} = The exponentially smoothed forecast made for the prior period
A_{t-1} = The actual demand in the prior period
α = The desired response rate, or smoothing constant

This equation states that the new forecast is equal to the old forecast plus a portion of the error (the difference between the previous forecast and what actually occurred).

Potential exercise in the exam

Example: the long-run demand for the product under study is relatively stable and a smoothing constant (α) of 0.05 is considered appropriate.
Last month's forecast (F_t–1) was 1,050 units.
If 1,000 actually were demanded, rather than 1,050, the forecast for this month would be:

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1}) = 1050 + 0.05(1000 - 1050) = 1050 + 0.05(-50) = 1047.5 \text{ units}$$

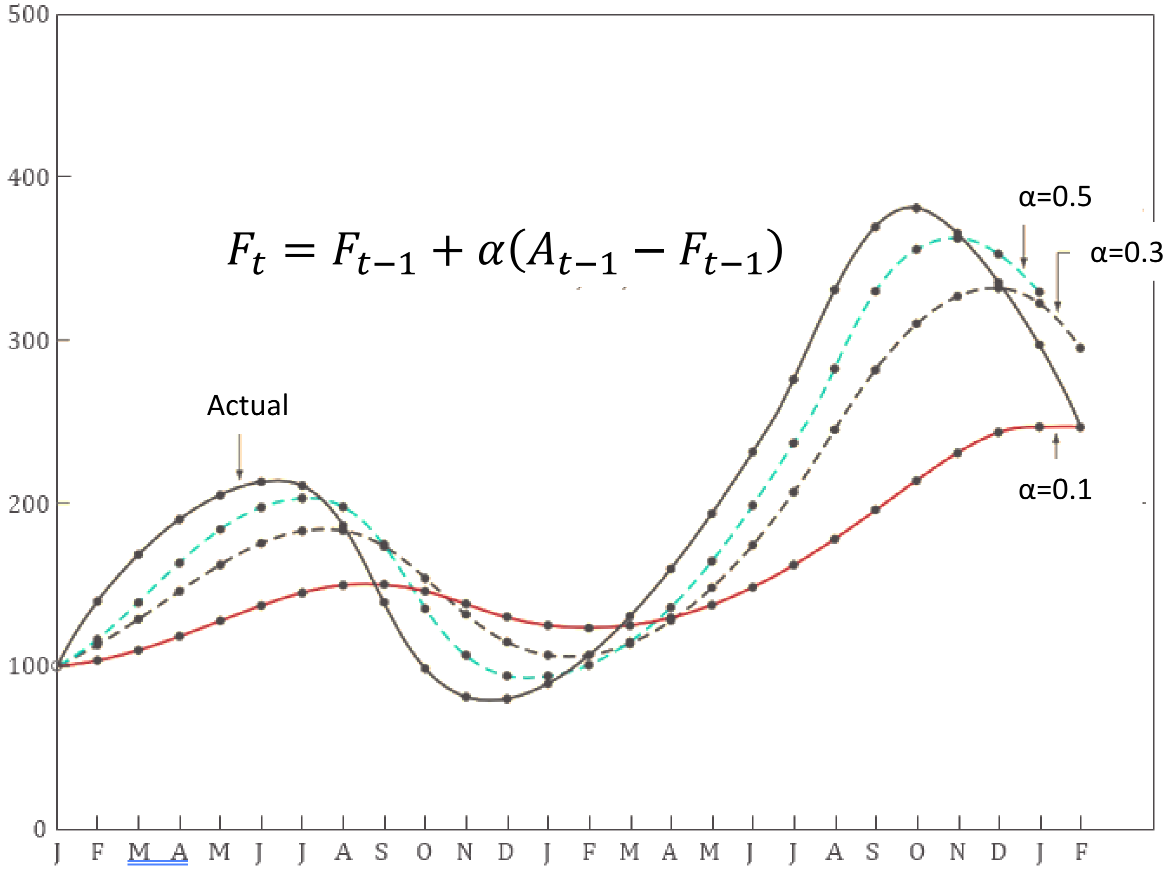
Because the smoothing coefficient is small, the reaction of the new forecast to an error of 50 units is to decrease the next month's forecast by only 2½ units.

Potential exercise in the exam

When exponential smoothing is first used for an item, an initial forecast may be obtained by using a simple estimate, like the first period's demand, or by using an average of preceding periods, such as the average of the first two or three periods.

Single exponential smoothing has the shortcoming of lagging changes in demand:

- The forecast lags during an increase or decrease, but overshoots when a change in direction occurs.
- The higher the value of alpha, the more closely the forecast follows the actual.
- To more closely track actual demand, a trend factor may be added. Adjusting the value of alpha also helps (adaptive forecasting).



Exponential Smoothing with Trend

Exponentially smoothed forecasts can be corrected somewhat by adding in a trend adjustment. To correct the trend, we two smoothing constants are needed:

- the smoothing constant α
- a smoothing constant delta (δ).

Both alpha and delta reduce the impact of the error that occurs between the actual and the forecast.
To get the trend equation going, the first time it is used the trend value must be entered manually (it can be an educated guess or a computation based on observed past data).
Equations to compute the forecast including trend (FIT):

$$1.\ F_t = FIT_{t-1} + \alpha(A_{t-1} - FIT_{t-1})$$
$$2.\ T_t = T_{t-1} + \delta(F_t - FIT_{t-1})$$

$$FIT_t = F_t + T_t$$

F_t = The exponentially smoothed forecast that does not include trend for period t T_t = The exponentially smoothed trend for period t
FIT_t = The forecast including trend for period t
FIT_{t-1} = The forecast including trend made for the prior period
A_{t-1} = The actual demand for the prior period
α = Smoothing constant (alpha)
δ = Smoothing constant (delta)

Steps:

- Step 1: Using Equation 1, make a forecast that is not adjusted for trend. This uses the previous forecast and previous actual demand.
- Step 2: Using Equation 2, update the estimate of trend using the previous trend estimate, the unadjusted forecast just made, and the previous forecast.
- Step 3: Make a new forecast that includes trend by using the results from steps 1 and 2.

Exponential smoothing requires that the smoothing constants be given a value between 0 and 1 (in the range of .1 to .3). The values depend on how much random variation there is in demand and how steady the trend factor is. Error measures can be helpful in picking appropriate values for these parameters.

Example: Assume a previous forecast, including a trend of 110 units, a previous trend estimate of 10 units, an alpha of .20, and a delta of .30. If actual demand turned out to be 115 rather than the forecast 110, the forecast for the next period is: calculated as:

- The actual A_t–1 is given as 115. Therefore,

$$F_t = FIT_{t-1} + \alpha(A_{t-1} - FIT_{t-1}) = 110 + 0.2(115 - 110) = 111.0$$
$$T_t = T_{t-1} + \delta(F_t - FIT_{t-1}) = 10 + 0.3(111 - 110) = 10.3$$
$$FIT_t = F_t + T_t = 111.0 + 10.3 = 121.3$$

- If, instead of 121.3, the actual turned out to be 120, the sequence would be repeated and the forecast for the next period would be:

$$F_{t+1} = 121.3 + 0.2(120 - 121.3) = 121.04$$
$$T_{t+1} = 10.3 + 0.3(121.04 - 121.3) = 10.22$$
$$FIT_{t+1} = 121.04 + 10.22 = 131.26$$

Linear Regression Analysis

A forecasting technique that fits a straight line to past demand data.

The linear regression line is of the form:

$$Y = a + bt$$

Potential exercise in the exam

Y is the value of the dependent variable that we are solving for

a is the Y intercept

b is the slope

t is an index for the time period.

Linear regression is useful for long-term forecasting of major occurrences and aggregate planning.

The major restriction in using linear regression forecasting is, as the name implies, that past data and future projections are assumed to fall in about a straight line (used for short period of time).

- When the dependent variable changes as a result of time, it is time series analysis.
- If one variable changes because of the change in another variable, this is a causal relationship.

Example:
A firm’s sales for a product line during the 12 quarters of the past three years:

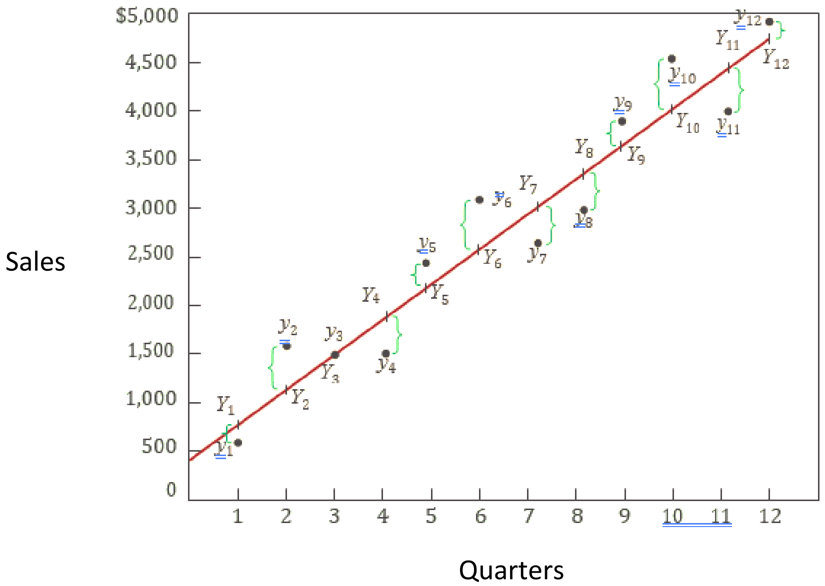
Quarter	Sales		Quarter	Sales
1	600		7	2,600
2	1,550		8	2,900
3	1,500		9	3,800
4	1,500		10	4,500
5	2,400		11	4,000
6	3,100		12	4,900

Potential exercise in the exam

The firm wants to forecast each quarter of the fourth year (quarters 13, 14, 15, and 16).

$$Y = a + bt$$

The least squares method tries to fit the line to the data that minimizes the sum of the squares of the vertical distance between each data point and its corresponding point on the line.
If a straight line is drawn through the general area of the points, the difference between the point and the line is y – Y.



The sum of the squares of the differences between the plotted data points and the line points is:

$$(y_1 - Y_1)^2 + (y_2 - Y_2)^2 + \cdots + (y_{12} - Y_{12})^2$$

The best line to use is the one that minimizes this total.

The equations for a and b are:

$$b = \frac{\sum ty - n\overline{ty}}{\sum t^2 - \bar{n}t^2}$$

$$a = \bar{y} - b\bar{t}$$

Potential exercise in the exam

a = Y intercept
b = Slope of the line
$\bar{y}$ = Average of all ys
$\bar{t}$ = Average of all ts
t = t value at each data point
y = y value at each data point
n = Number of data points
Y = Value of the dependent variable computed with the regression equation

Potential exercise in the exam

(1) t	(2) y	(3) t × y	(4) t^2	(5) y^2	(6) y
1	600	600	1	360,000	801.3
2	1,550	3,100	4	2,402,500	1,160.9
3	1,500	4,500	9	2,250,000	1,520.5
4	1,500	6,000	16	2,250,000	1,880.1
5	2,400	12,000	25	5,760,000	2,239.7
6	3,100	18,600	36	9,610,000	2,599.4
7	2,600	18,200	49	6,760,000	2,959.0
8	2,900	23,200	64	8,410,000	3,318.6
9	3,800	34,200	81	14,440,000	3,678.2
10	4,500	45,000	100	20,250,000	4,037.8
11	4,000	44,000	121	16,000,000	4,397.4
12	4,900	58,800	144	24,010,000	4,757.1
78	33,350	268,200	650	112,502,500	
$\bar{t}$ = 6.5 b = 359.6154					
$\bar{y}$ = 2,779.17 a = 441.6667					
Therefore, Y = 441.67 + 359.6t					
S_{yt} = 363.9					

The slope shows that for every unit change in t, Y changes by 359.6.
Strictly based on the equation, forecasts for periods 13 through 16 would be

$$Y_{13} = 441.67 + 359.6(13) = 5116.5$$

$$Y_{14} = 441.67 + 359.6(14) = 5476.1$$

$$Y_{15} = 441.67 + 359.6(15) = 5835.7$$

$$Y_{16} = 441.67 + 359.6(16) = 6195.3$$

The standard error of estimate, or how well the line fits the data, is:

$$S_{yt} = \frac{\sqrt{\sum_{i=1}^{n}(y_i - Y_i)^2}}{n-2}$$

$$S_{yt} = \frac{\sqrt{(600-801.3)^2 + \cdots + (4900-4757.1)^2}}{10} = 363.9$$

Decomposition of a Time Series

The process of identifying and separating time series data into fundamental components such as trend and seasonality.

Two types of seasonal variation:

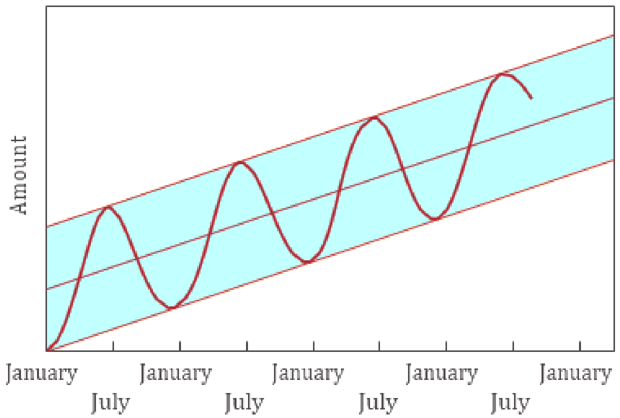
- Additive seasonal variation simply assumes that the seasonal amount is a constant no matter what the trend or average amount is:

Forecast including trend and seasonal = Trend + Seasonal

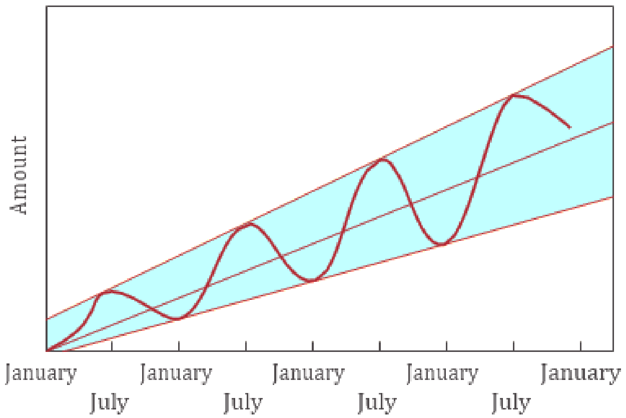
- In multiplicative seasonal variation, the trend is multiplied by the seasonal factors.

Forecast including trend and seasonal = Trend × Seasonal factor

Additive seasonal



Multiplicative seasonal



Example:

In past years, a firm sold an average of 1,000 units of a particular product line each year. On the average, 200 units were sold in the spring, 350 in the summer, 300 in the fall, and 150 in the winter.

The seasonal factor (or index) is the ratio of the amount sold during each season divided by the average for all seasons.

The yearly amount divided equally over all seasons is 1,000 ÷ 4 = 250.

The seasonal factors therefore are:

	Past Sales	Average Sales for Each (1,100/4) Season	Seasonal Factor
Spring	200	250	200/250 = 0.8
Summer	350	250	350/250 = 1.4
Fall	300	250	300/250 = 1.2
Winter	150	250	150/250 = 0.6
Total	1,000		

Using these factors, if the demand for next year is expected to be 1,100 units, the demand is forecasted as:

	Expected Demand for Next Year	Average Sales for Each (1,100/4) Season		Seasonal Factor		Next Year's Seasonal Forecast
Spring		275	×	0.8	=	220
Summer		275	×	1.4	=	385
Fall		275	×	1.2	=	330
Winter	____	275	×	0.6	=	165
Total	1,100					

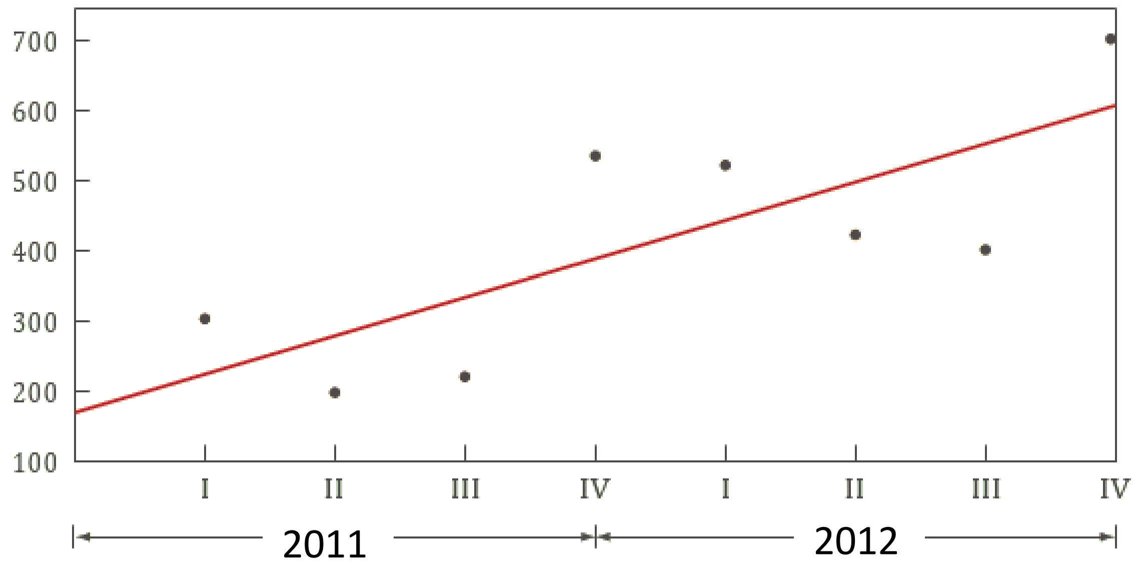
The seasonal factor may be periodically updated as new data are available.

Example:

Forecast the demand for each quarter of the next year using trend and seasonal factors.

Demand for the past two years is:

Quarter	Amount		Quarter	Amount
1	300		5	520
2	200		6	420
3	220		7	400
4	530		8	700



$$Forecast\ Including\ Trend\ (FIT) = 176.1 + 52.3t$$

Next, a seasonal index can be derived by comparing the actual data with the trend line.

The seasonal factor was developed by averaging the same quarters in each year.

Quarter	Actual amount	From trend equation	Ratio of actual+trend	Seasonal factor (avg. Of the same quarters in both years)
2011				
I	300	228.3	1.31	
II	200	280.6	0.71	
III	220	332.9	0.66	I 1.25
IV	530	385.1	1.38	II 0.79
2012				III 0.70
I	520	437.4	1.19	IV 1.28
II	420	489.6	0.86	
III	400	541.9	0.74	
IV	700	594.2	1.18	

The 2013 forecast is computed including trend and seasonal factors (FITS): $FITS_t = FIT\ x\ seasonal$

$$I - 2013FITS_9 = \big(176.1 + 52.3(9)\big)1.25 = 808$$

$$II - 2013FITS_{10} = \big(176.1 + 52.3(10)\big)0.79 = 502$$

$$III - 2013FITS_{11} = \big(176.1 + 52.3(11)\big)0.70 = 526$$

$$IV - 2013FITS_{12} = \big(176.1 + 52.3(12)\big)1.28 = 1029$$

Forecast Errors

The difference between actual demand and what was forecast.

Sources of Error Errors can come from a variety of sources, e.g. projecting past trends into the future. Errors can be classified as:

- Bias errors occur when a consistent mistake is made.
- Random errors can be defined as those that cannot be explained by the forecast model being used.

Measurement of Error

- Mean absolute deviation (MAD) is the average of the absolute value of the actual forecast error. It measures the dispersion of some observed value from some expected value. MAD is computed using the differences between the actual demand and the forecast demand without regard to sign:

$$MAD = \frac{\sum_{t=1}^{n} |A_t - F_t|}{n}$$

t = Period number
A_t = Actual demand for the period t F_t = Forecast demand for the period t n = Total number of periods
|| = A symbol used to indicate the absolute value disregarding positive and negative signs

Potential exercise in the exam

When the errors that occur in the forecast are normally distributed (the usual case), the mean absolute deviation relates to the standard deviation of the error terms:

$$a\ standard\ deviation \cong MAD\sqrt{\frac{\pi}{2}}\ or\ 1.25MAD$$

- Mean absolute percent error (MAPE) is the average error measured as a percentage of average demand.

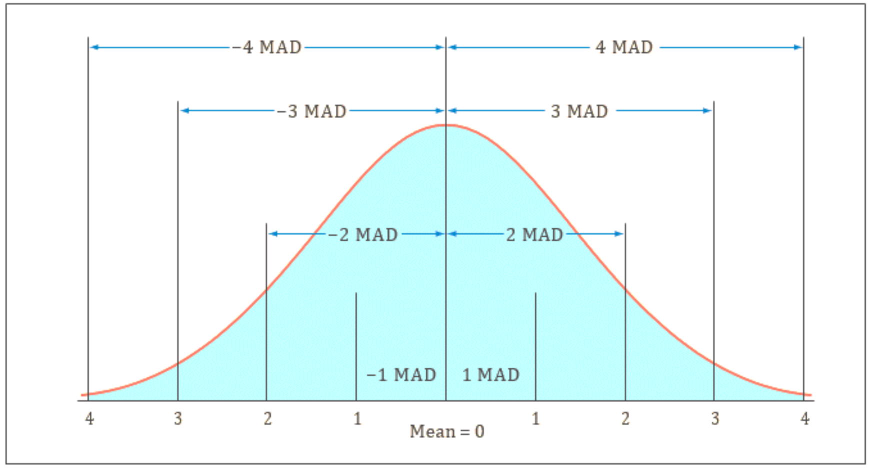
$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{|A_t - F_t|}{A_t}$$

The real value of the MAPE is that it allows to compare forecasts between products that have very different average demand.

Potential exercise in the exam

- Tracking signal is a measure of whether the forecast is keeping pace with any genuine upward or downward changes in demand. This is used to detect forecast bias.

$$TS = \frac{RSFE}{MAD}$$



RSFE = The running sum of forecast errors, considering the nature of the error.

Month	Demand Forecast	Actual	Deviation	RSFE	Abs. Dev.	Sum of Abs. Dev.	MAD (% Error)*	TS = RSFE† / MAD
1	1,000	950	−50	−50	50	50	50 (5.26%)	−1
2	1,000	1,070	+70	+20	70	120	60 (5.61%)	.33
3	1,000	1,100	+100	+120	100	220	73.3 (6.67%)	1.64
4	1,000	960	−40	180	40	260	65 (6.77%)	1.2
5	1,000	1,090	+90	+170	90	350	70 (6.42%)	2.4
6	1,000	1,050	+50	+220	50	400	66.7 (6.35%)	3.3

*Overall, MAD = 400 ÷ 6 = 66.7. MAPE = (5.26 + 5.61 + 6.67 + 6.77 + 6.42 + 6.35)/6 = 6.18%

†Overall, TS = $\frac{RSFE}{MAD} = \frac{220}{66.7}$ = 3.3 MADs.

Potential exercise in the exam

If the actual demand does not fall below the forecast to offset the continual positive RSFE, the tracking signal would continue to rise and we would conclude that assuming a demand of 1,000 is a bad forecast.

Qualitative forecasting techniques

They generally take advantage of the knowledge of experts and require much judgment. These techniques typically involve processes that are well defined to those participating in the forecasting exercise:

- Market research. It is used mostly for product research in the sense of looking for new product ideas, likes and dislikes about existing products, which competitive products within a particular class are preferred, and so on. The data collection methods are primarily surveys and interviews.
- Panel Consensus. It is based on the idea that a panel of people from a variety of positions can develop a more reliable forecast than a narrower group. Panel forecasts are developed through open meetings with a free exchange of ideas from all levels of management and individuals.
- Historical Analogy. An existing product or generic product is used as a model to forecast demand for a new product
- Delphi Method
    1. Choose the experts to participate. There should be a variety of knowledgeable people in different areas.
    2. Through a questionnaire (or e-mail), obtain forecasts (and any premises or qualifications for the forecasts) from all participants.
    3. Summarize the results, and redistribute them to the participants along with appropriate new questions.
    4. Summarize again, refining forecasts and conditions, and again develop new questions.
    5. Repeat step 4 if necessary. Distribute the final results to all participants.

The role of inventory management is to strike a balance between inventory investment and customer service.

Examples of inventory functions:

- to decouple the company and the supply chain from fluctuations in demand and hold a stock of goods that will provide a selection for customers;
- to increase supply chain flexibility by placing inventory in the right places;
- to hedge against facility disruptions in the event of natural catastrophes;
- to decouple or separate various parts of the production process;
- to take advantage of quantity discounts and hedge against inflation.

Inventory is classified according to the following types:

- Raw material: items which are purchased but not processed;
- Work-in-process (WIP): items which underwent some changes, but are not completed;
- Maintenance/repair/operating (MRO): items which are necessary to keep machinery and processes productive;
- Finished goods: completed product awaiting shipment.

In calculating inventory amounts, the following costs are typically considered:

- Holding costs (variable): the costs of holding inventory over time;
- Ordering costs (fixed): the costs of placing an order and receiving goods;
- Setup costs (fixed): the costs pf preparing a machine or process for manufacturing an order;
- Stockout costs (variable): the costs of lost customer orders resulting from product shortage, loss-of-goodwill costs.

According to inventory functions and types, inventory can be used to manage:

- Economy of scale. (cycle inventory). Cycle inventory exists as a result of producing or purchasing in large lots or batches (the quantity that a stage in the supply chain either produces or purchases at a time). The supply chain can exploit economy of scale and order in large lots to reduce fixed costs.
- Uncertainty (safety inventory). Safety inventory is carried to satisfy demand subject to unpredictable demand fluctuations and to reduce product shortages. Choosing safety inventory involves making a trade-off between the costs of having too much inventory and the costs of losing sales due to inventory shortage.

Seasonal inventory is built up to counter predictable variability in demand. Companies using seasonal inventory build up inventory in periods of low demand and store it for periods of high demand when they will not have the capacity to produce all that is demanded.

Material analysis

ABC analysis

ABC analysis divides inventory into three classes based on annual dollar volume:

- Class A: high annual dollar volume
- Class B: medium annual dollar volume
- Class C: low annual dollar volume

ABC analysis is used to establish policies that focus on the few critical parts and not the many trivial ones.

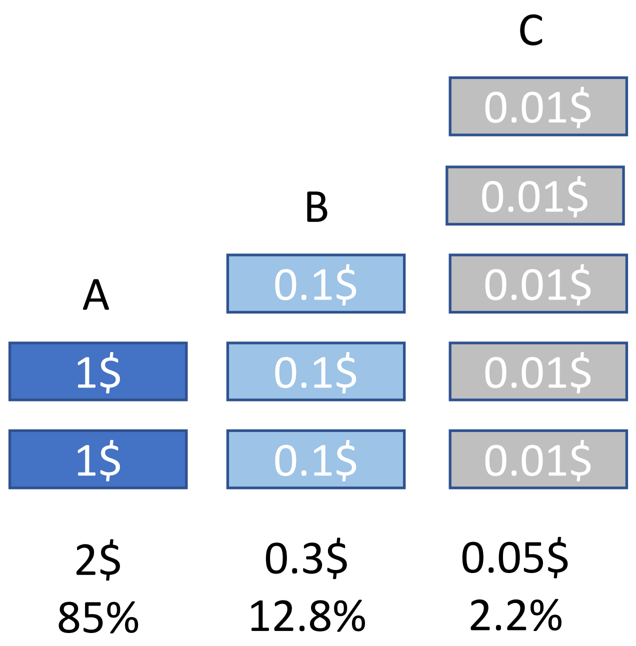
Potential exercise in the exam

Example:

- two red blocks (1$ each)
- three green blocks (0.1$ each)
- five blue blocks (0.01$ each)

Total inventory costs: $2.35.

Two red blocks take only 20% of the total inventory amount, but they create 85% of inventory costs. These are critical items in the A group.



Example:
A company which sells table lamps.
To determine which table lamps are in the A, B, or C-category:

Potential exercise in the exam

- the annual expenditure is estimated:

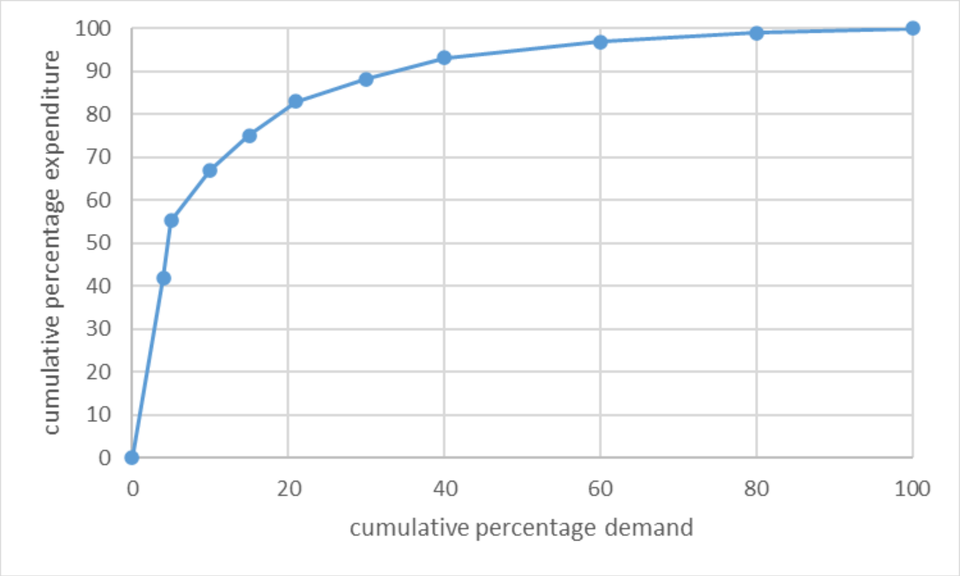
$$Annual\ expenditure = annual\ demand - cost\ per\ lamp$$

- Cumulative expenditure and the corresponding percentage is calculated
- different types of table lamps are classified as A, B, and C items according to their annual expenditure, assuming that the classification is 80:15:5

Table lamp	Annual demand	Cost per unit	Annual expenditure
X1	100	0.5	50
X2	200	0.05	10
X3	50	1.65	82.5
Y1	40	10.75	430
Y2	200	0.11	22
Y3	200	0.19	38
Y4	50	2.4	120
T1	90	0.6	54
T2	10	13.6	136
T3	60	1.35	81

Table lamp	Annual demand	Cost per unit	Annual expenditure	Cumulative expenditure	Percentage expenditure	Category
Y1	40	10.75	430	430	42.0	A
T2	10	13.6	136	566	55.3	A
Y4	50	2.4	120	686	67.0	A
X3	50	1.65	82.5	768.5	75.1	A
T3	60	1.35	81	849.5	83.0	B
T1	90	0.6	54	903.5	88.3	B
X1	100	0.5	50	953.5	93.2	B
Y3	200	0.19	38	991.5	96.9	C
Y2	200	0.11	22	1013.5	99.0	C
X2	200	0.05	10	1023.5	100.0	C

150 table lamps generate around 80% of inventory costs (A), whereas 600 table lamps generate only 5% of inventory costs (C).



Potential exercise in the exam

It can be observed that 80% of expenditure is created by only 20% of demand.

XYZ Analysis

Besides the monetary value of capital commitment, other criteria can be used to classify items (XYZ analysis). One option is to divide inventory into three classes based on different demand patterns:

- Class X: constant, non-changing demand;
- Class Y: neither constant nor sporadic demand (fluctuating demand);
- Class Z: sporadic or strongly fluctuating demand.

Changes in demand make it possible to determine the prediction accuracy of each inventory class.
XYZ analysis can be used to enhance ABC analysis.

Annual dollar volume

Changes in demand

	A	B	C
X	High volume Constant demand	Middle volume Constant demand	Low volume Constant demand
Y	High volume Fluctuating demand	Middle volume Fluctuating demand	Low volume Fluctuating demand
Z	High volume Sporadic demand	Middle volume Sporadic demand	Low volume Sporadic demand
	Upon need	Just in time	Demand-oriented

Example:
Store that needs to reduce the space in its warehouse by 50% and costs by 70%.

	Quantity	Cost perunit ($)	Annual expenditure	Volume per unit (dm^3)	Total volume (dm^3)
Blouses	120	200	24,000	1.00	120
Pantsuits	420	200	84,000	1.00	420
Jeans	50	200	10,000	2.00	100
Dresses and skirts	450	500	225,000	2.00	900
Costumes	280	1000	280,000	2.00	560
Fur coats	120	10,000	1,200,000	10.00	1200
Sport pants	10	80	800	2.00	20
T-shirts	1200	80	96,000	0.50	600
Scarfs	100	50	5000	0.40	40
Underwear	500	75	37,500	0.02	10
Belts	600	95	57,000	0.05	30
Total	3850		2,019,300		4000

	Quantity	Frequentuse	Seldomuse	No use	% of no use from quantity
Blouses	120	30	30	60	50.00
Pantsuits	420	100	120	200	47.62
Jeans	50	20		30	60.00
Dresses andskirts	450	10	20	420	93.33
Costumes	280	20	30	230	82.14
Fur coats	120	5	15	100	83.33
Sport pants	10	4	2	4	40.00
T-shirts	1200	300	400	500	41.67
Scarfs	100	30	30	40	40.00
Underwear	500	200	100	200	40.00
Belts	600	200	200	200	33.33

- items are sorted according to their volume

	Quantity	Volume per unit(dm³)	Total volume(dm³)	Percentage	Cumulativepercentage
Fur coats	120	10	1200	30	30
Dresses andskirts	450	2	900	22.5	52.5
T-shirts	1200	0.5	600	15	67.5
Costumes	280	2	560	14	81.5
Pantsuits	420	1	420	10.5	92
Blouses	120	1	120	3	95
Jeans	50	2	100	2.5	97.5
Scarfs	100	0.4	40	1	98.5
Belts	600	0.05	30	0.75	99.25
Sport pants	10	2	20	0.5	99.75
Underwear	500	0.02	10	0.25	1
Total	3850	20.97	4000	100.00	100

- ABC analysis is performed

	Quantity	Cost perunit ($)	Annual expenditure	Cumulative expenditure	Percentage expenditure	Category
Fur coats	120	10,000	1,200,000	1,200,000	59.4	A
Costumes	280	1000	280,000	1,480,000	73.3	A
Dresses and skirts	450	500	225,000	1,705,000	84.4	B
T-shirts	1200	80	96,000	1,801,000	89.2	B
Pantsuits	420	200	84,000	1,885,000	93.4	B
Belts	600	95	57,000	1,942,000	96.2	C
Underwear	500	75	37,500	1,979,500	98.0	C
Blouses	120	200	24,000	2,003,500	99.2	C
Jeans	50	200	10,000	2,013,500	99.6	C
Scarfs	100	50	5000	2,018500	99.7	C
Sportpants	10	80	800	2,019,300	100	C
Total	3850		2,019,300	2,019,300	100	

- items are sorted according to the XYZ classification in order to see how much potential there is for space reduction (X-items and the percentage of never used X-items, which are primary candidates for leaving the wardrobe are identified).
- In parallel, A-items are identified, since the second goal in this task is also to reduce expenditure to 70%

	Quantity	% of neverused items	Cumulative percentage(volume)	Space saving(%)	ABC	Cost saving
Fur coats	120	50.00	30	25.00	A	1000,000
Dresses and skirts	450	47.62	52.5	21.00	B	210,000
T-shirts	1200	60.00	67.5	6.25	B	40,000
Costumes	280	93.33	81.5	11.50	A	230,000
Pantsuits	420	82.14	92	5.00	B	40,000
Blouses	120	83.33	95	1.50	C	12,000
Jeans	50	40.00	97.5	1.50	C	6000
Scarfs	100	41.67	98.5	0.40	C	2000
Belts	600	40.00	99.25	0.25	C	19,000
Sportpants	10	40.00	99.75	0.20	C	3200
Underwear	500	33.33	100	0.10	C	15,000
Total	3850		100	100		1,577,200

It can be observed that due to the high value, high volume, and high percentage of never used items (the fur coats, dresses, skirts, and costumes) are key in achieving both the objectives, i.e. value reduction of 70% and space reduction of 50%.

Deterministic Models

These models determine order quantities considering items with independent deterministic demand and lead time (i.e. (time from ordering to receipt) with the aim to exploit economy of scale and order in large lots to reduce fixed ordering costs.

The following models are presented:

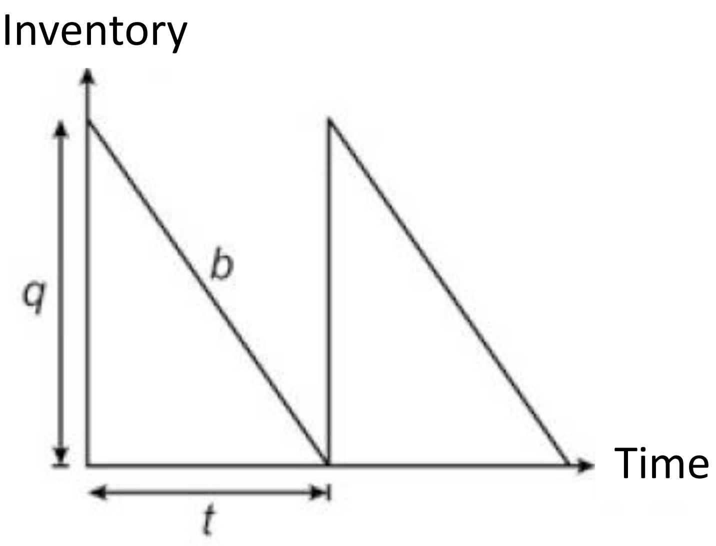
- Basic economic order quantity (EOQ)
- Quantity discount model
- Economic production order quantity (EPQ)
- Reorder point (ROP).

Potential exercise in the exam

EOQ Model

Consider the system that exhibits the following characteristics:

- Demand and lead-time are known and constant;
- Receipt of inventory is instantaneous and complete;
- Quantity discounts are not possible;
- The only variable costs are setup and holding;
- Stock-out can be avoided.



- q is the number of units per order;
- q* is optimal number of units per order (EOQ);
- b is annual demand in units for the inventory item;
- f is set-up or ordering cost for each order;
- c is holding or carrying cost per unit per year.

Under the assumption of linear inventory consumption, cycle inventory, and lot-sizes are related as:
$Cycle\ inventory = {}^{q}/_{2}$

The annual inventory holding costs is : ${}^{cq}/_{2}$

The number of orders per year: ${}^{b}/_{q}$

The annual fixed ordering costs is: ${}^{fb}/_{q}$

Potential exercise in the exam

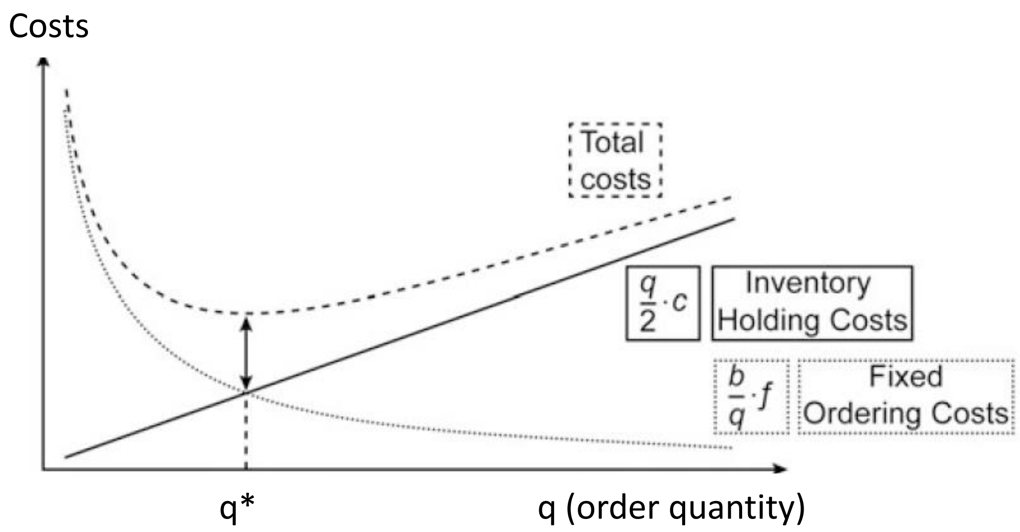
Optimal order quantity is found when annual ordering costs equal annual holding costs:

$$\frac{cq}{2} = \frac{fb}{q^*}$$

The EOQ formula is obtained by solving the equation for q*:

$$q^* = \sqrt{\frac{2bf}{c}}$$

Potential exercise in the exam



- the smallest total cost (the top curve) is the sum of the two curves below it.
- minimal total costs are achieved at the intersection point of the fixed and variable costs curves (EOQ point q*).
- total cost function is quite flat in the minimal region. This means that moderate EOQ changes will not influence a significant increase in total costs by tendency.

The expected number of orders per year is:

$$N = \frac{b}{q^*}$$

The expected time between orders is:

$$T = \frac{365}{N}$$

The total annual cost is:

$$TC = \frac{cq^*}{2} + \frac{fb}{q^*}$$

Potential exercise in the exam

Example:

- Demand for a TV at a store is 3200 units per quarter.
- The store charges fixed costs of $2500 per order.
- Annual holding costs per TV are $80.

Calculate the number of TVs that the store manager should order per refill.

- The Annual demand is b=3200 x 4 = 12800 units;
- The ordering cost per order is f=$2500
- The holding cost per unit per year is: c=$80
- Using the EOQ formula, the optimal order quantity is: $q^* = \sqrt{\frac{2 \cdot 12800 \cdot 2500}{80}} = 895$
- The cycle inventory is: 895/2=448 units
- Number of orders per year: N = 12,800/895 = 14.3
- Expected time between orders: T = 365/14.3 = 25.5 days
- Total cost: TC = 80 x 895/2 + 2500 x 12,800/895 = $71,554
- Total cost for 1 year is $71,554 for 14.3 orders with 895 TVs in each order.

Potential exercise in the exam

EOQ is not optimal for the integrated inventory-transportation setting.

Example:
EOQ is 40 units. This corresponds to 25 deliveries a year subject to annual demand of 1000 modules. Consider the following procedure for determining transportation costs: 400 € per delivery + 4 € per module. It is possible to transport up to 100 units at a time.

- Costs Analysis for 100 Units:
  - Transportation costs: 10 deliveries x (400 + 4 x 100)=8 € per unit
  - Cycle inventory: (100/2) x 29 + (1000/100) x 23.2=1682 €
  - Safety inventory: 1.65 x 4 x 10=66 items x 29=1914 €
  - Total costs per module: 11.6 €
- Costs Analysis for 40 Units:
  - Transportation costs: 25 deliveries x (400 + 4 x 40)=14 € per unit
  - Cycle inventory: (40/2) x 29 + (1000/40) x 23.2=1160 €
  - Safety inventory: 1.65 x 4 x 10=66 units x 29=1914 €
  - Total costs per module: 17.1 €

Potential exercise in the exam

EOQ Model with Discounts

Inventory costs can also be calculated on the basis of the unit prices p (i.e., the actual costs of the material purchased) as:

$$Cost(q) = p_1(1 - r_1)b + \frac{q}{2}p_1(1 - r_1)I + \frac{bf}{q}$$

For calculating EOQ:

$$q^* = \sqrt{\frac{2bf}{pI}}$$

where I is the interest rate (capital commitment) and p is the unit price.

The formula allows to apply the EOQ model for different prices (e.g. situations with quantity discounts). Reduced prices are often available when larger quantities are purchased. In this case, the trade-off is between reduced item costs and increased holding costs.

The algorithm of calculating EOQ with discounts involves the following steps:

1. For each discount, calculate q*;
2. If q* does not qualify for a discount, choose the smallest possible order size to get the discount;
3. Compute the total cost for each q* or adjusted value from Step 2;
4. Select the q* that gives the lowest total cost.

Example:

A chocolate shop where:

- Cost for one unit of chocolate is 5.00$, but a quantity discount is provided by the manufacturer:

Discount quantity in units	Discount (%)	Discount price p ($)
0–999	0	5.00
1000–1999	4	4.80
2000–10,000	10	4.50

- Annual demand for chocolate is 10,000 units
- The setup cost per order is 50$. Interest rate is 20%.

- For discount 0%: $q^* = \sqrt{\frac{2 \cdot 10000 \cdot 50}{5 \cdot 0.2}} = 1000$ and $Cost(1000) = 5 \cdot (1-0) \cdot 10000 + \frac{1000}{2} \cdot 5 \cdot (1-0) \cdot 0.2 + \frac{10000}{1000} \cdot 50 = 51000\$$
- For discount 4%: $q^* = \sqrt{\frac{2 \cdot 10000 \cdot 50}{4.8 \cdot 0.2}} = 1021$ and (q*∈ [1000-1999], therefore it can be used to compute the cost) $Cost(1000) = 5 \cdot (1-0.04) \cdot 10000 + \frac{1021}{2} \cdot 5 \cdot (1-0.04) \cdot 0.2 + \frac{10000}{1021} \cdot 50 = 48980\$$
- For discount 10%: $q^* = \sqrt{\frac{2 \cdot 10000 \cdot 50}{4.58 \cdot 0.2}} = 1054$ and (q*∋ [2000-10000], therefore the smallest possible order size to get the discount of 10% is used for the cost) $Cost(1000) = 5 \cdot (1-0.1) \cdot 10000 + \frac{2000}{2} \cdot 5 \cdot (1-0.1) \cdot 0.2 + \frac{10000}{2000} \cdot 50 = 46150\$$
- N=10000/2000=5
- T=365/5=73

EPQ Model

It is applied to manufacturing and it is used when:

- inventory builds up over a period of time after an order is placed;
- units are produced and sold simultaneously.

The receipt of inventory is allowed over a period of time.

Holding costs are calculated subject to the relation of production and demand:

$$\frac{cq}{2}\left(1-\frac{d}{r}\right)$$

- r is daily production rate;
- d is daily demand;
- t is the length of the production run in days.

The EPQ formula is:

$$q^* = \sqrt{\frac{2bf}{\left(1-\frac{d}{r}\right)c}}$$

Maximal inventory level in the system is:

$$I^{max} = q\left(1-\frac{b}{r}\right)$$

Example:
A firm produces high-quality food processors. It sells 18,000 processors per year and is able to produce 125 machines per day. It works 250 days per year. Annual holding costs per food processor is $18 and setup costs are $800.
The economic production quantity for the firm and the maximal inventory level can be estimated as:

$$d = \frac{18000}{250} = 72 \; units \; per \; day$$

$$q^* = \sqrt{\frac{2 \cdot 18000 \cdot 800}{\left(1 - \frac{72}{125}\right) \cdot 18}} = 1943$$

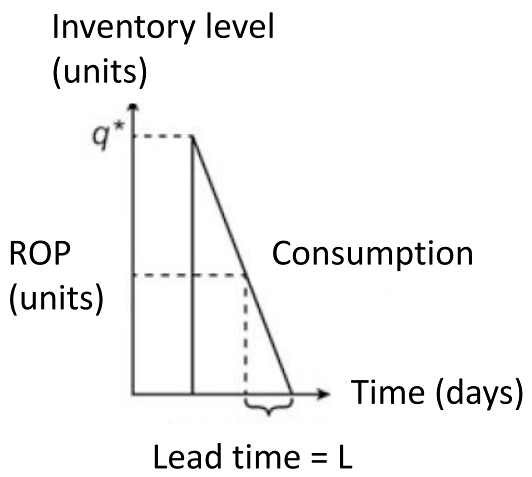
$$I_{max} = 1943 \cdot \left(1 - \frac{72}{125}\right) = 824$$

Re-order Point

The EOQ model answers the “how much” question. The re-order point (ROP) tells “when” to order. ROP is introduced to take into account the lead time, i.e. the time between placement and receipt of an order.
With the assumption of constant demand and a set lead time, ROP is calculated as:

$$ROP = dL$$

where d is daily demand and L is lead time.



Example:
A company experiences an annual demand of 8500 knives per year (250 working days). Lead time for an order is 5 working days.

$$Daily\ demand = \frac{8500}{250} = 34\ units$$
$$ROP = 34 \cdot 5 = 170\ units$$

Potential exercise in the exam

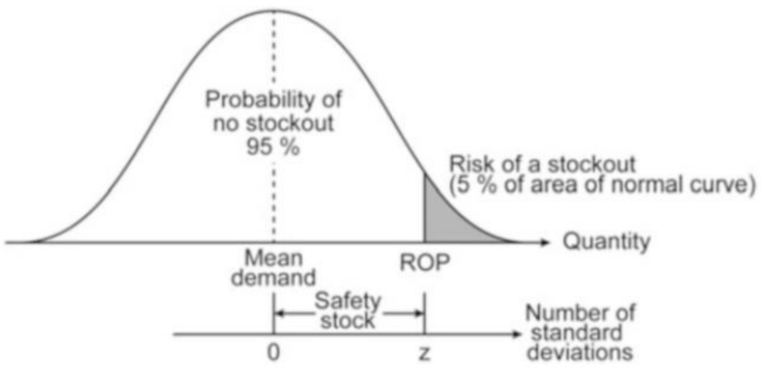
Stochastic Models

In many practical cases, both demand and lead time fluctuate. Their values are not known, but they can be estimated on the basis of probability.

Uncertainty in demand makes it necessary to maintain a certain customer service level or level of product availability (the fraction of demand that is served on time from a product held in inventory) to avoid stock-outs:

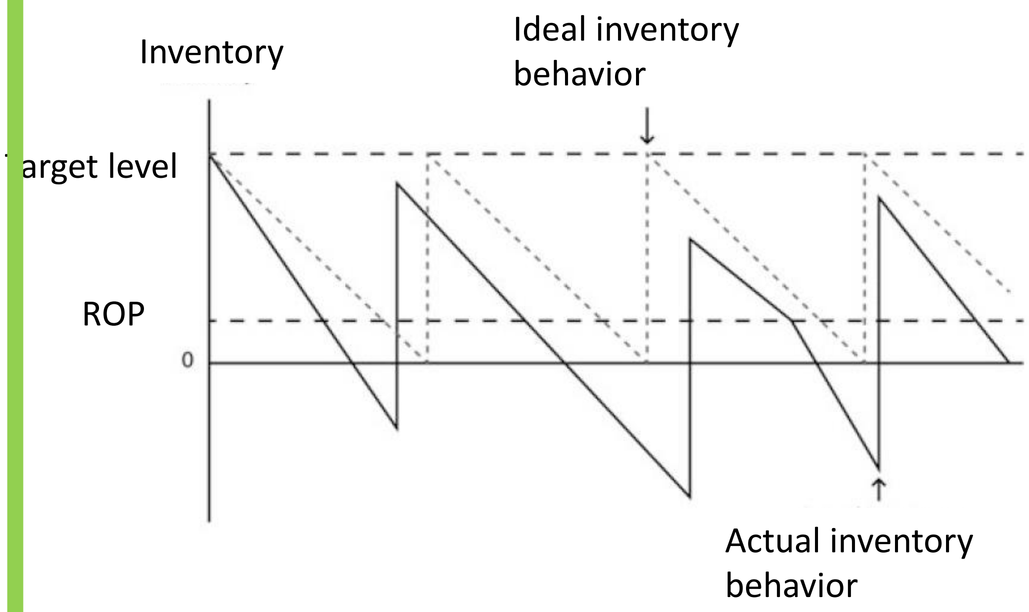
- A high level of product availability provides a high level of responsiveness, but increases costs because much inventory is held, but rarely used.
- A low level of product availability lowers inventory holding cost, but results in a higher fraction of customers who are not served on time.

The basic trade-off when determining the level of product availability is between the cost of inventory to increase product availability and the loss from not serving customers on time.

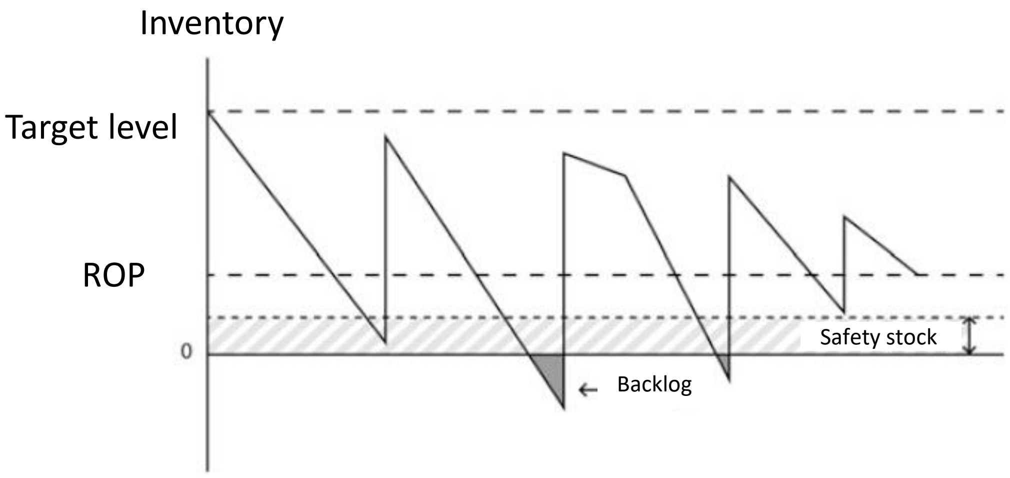


For example, a 0.05 probability of stock-out corresponds to a 95% service level. The higher the service level, the higher supply chain responsiveness, but also the higher inventory costs.

In a situation of demand uncertainty, safety inventory is introduced with the objective to ensure product availability even in the case of demand fluctuations.

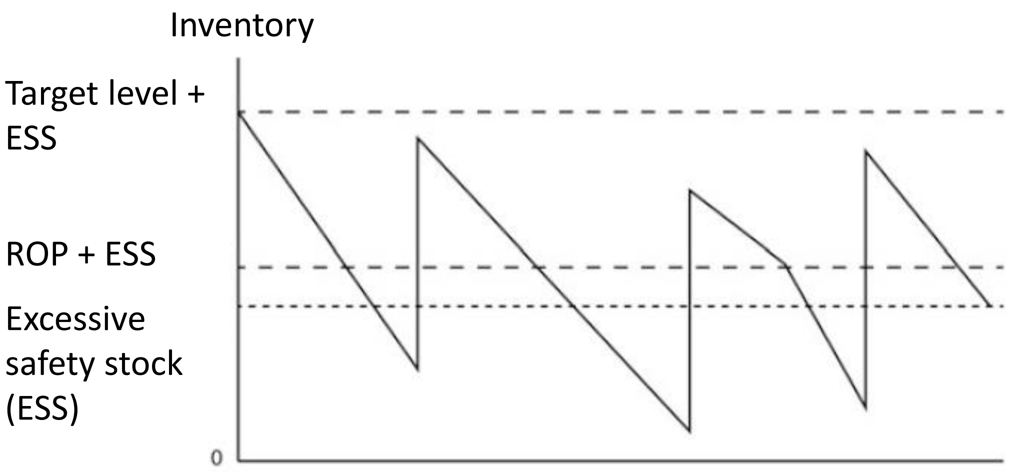


- The ideal inventory behavior means in this case that all assumptions of EOQ and ROP models are met, i.e., demand and lead-time are constant.
- In reality, this is not the case. Both demand and lead-time fluctuate resulting in actual inventory behavior which is different as the ideal one.
- In order to cope with this situation, the ROP should be increased by the safety stock.



ROP with safety stock and backlogs

The ROP is increased by safety stock.
Safety stock allows to cope with demand fluctuations in some cases and, in other cases there exists a backlog (the demand not satisfied in the period in which it arises, but carried over to future periods).



ROP with excessive safety stock and without backlogs

The ROP from is increased again by an excessive safety stock (ESS). The ESS is so high that demand fluctuations would never result in a backlog (a 100% product availability on stock resulting in a 100% service level).

However, the inventory is much higher resulting in higher inventory costs.

To compute safety stock subject to a desired service level:

$$ss = z\sigma_{dLT}$$

Where:

- ss is safety stock
- σ_{dLT} is standard deviation of demand during lead-time. It can be obtained from analysis of demand forecasts and actual sales in the past (σ=1.25MAD is a typical value).
- z is the number of standard deviations. It can be determined from the table of normal distribution.

Potential exercise in the exam

Example:
If standard deviation of demand during lead time is 10:

- for a service level of 99%, z=2.33, ss=23.3
- for a service level of 95%, z=1.65, ss=16.5
- for a service level of 90%, z=1.28, ss=12.8

The ROP is:

$$ROP = \bar{d}L + ss$$

To calculate ROP, four situations are possible:

- demand is assumed to be normally distributed during the lead time:

$$ROP = \bar{d}L + z\sigma_{dLT}$$

- daily distribution of demand is given (i.e., demand is variable) and lead time is constant:

$$ROP = \bar{d}L + z\sigma_d\sqrt{L}$$

- daily demand is constant and lead time is variable:

$$ROP = \bar{d}L + z\bar{d}\sigma_L$$

- both demand and lead time are variable:

$$ROP = \bar{d}L + z\sqrt{L\sigma_d^2 + \bar{d}^2\sigma_L^2}$$

Potential exercise in the exam

Example:

Average demand for toothbrushes is 35 units per day. Standard deviation of normally distributed demand during lead time is ten toothbrushes per day. Lead time is 3 days. Service level is 95%.

$$ROP = 35 \cdot 3 + 1.65 \cdot 10 = 122\, units$$

$$ss = 1.65 \cdot 10 = 16.5\, units$$

If a daily distribution of demand is given instead of standard deviation of normally distributed demand during lead time. And daily standard deviation of demand is 10 units:

$$ROP = 35 \cdot 3 + 1.65 \cdot 10 \cdot \sqrt{3} = 134\, units$$

If demand is constant, but lead time may fluctuate with a standard deviation of 1 day:

$$ROP = 35 \cdot 3 + 1.65 \cdot 35 \cdot 1 = 163\, units$$

Potential exercise in the exam

If both demand and lead time are variable:

$$ROP = 35 \cdot 3 + 1.65 \cdot \sqrt{3 \cdot 10^2 + 35^2 \cdot 1^2} = 170\, units$$

The newsvendor problem is a mathematical model for calculating the optimal inventory level for one single period. A newspaper vendor who must decide every day how many daily newspapers he wants to stock for the next day is faced by uncertain demand and the knowledge that unsold copies will be almost worthless next day.
It is characterized by the following conditions:

- fixed price for each unit,
- perishable product,
- uncertain demand,
- no additional delivery in period t,
- short purchase time.

The following notation is used:
c is purchase price; r is retail price; v is salvage price;
c_o is overage cost; c_u is underage cost;
z is the number of standard deviations; σ is the standard deviation of demand; μ is the expectation of demand;
S is order quantity;
S* is optimal order quantity;
Z(S*) is the expected cost for optimal order quantity;
∏(S*) is the expected profit for optimal order quantity.

To calculate the overage and underage costs:

$$c_o = c - v$$

$$c_u = r - c$$

The critical ratio (CR) is used to find the z-value from the table of normal distribution:

$$CR = \frac{c_u}{c_u + c_o} \rightarrow F(CR) = z$$

To calculate S*:

$$S^* = \mu + z\sigma$$

The expected cost is:

$$Z(S^*) = (c_o + c_u) f_{01}(z)\sigma$$

(f_{01}(z) value can be taken from the full version of a normal distribution table).

The profit for optimal order quantity is:

$$\Pi(S^*) = c_u(\mu - Z(S^*))$$

Example:
A coffee shop purchases croissants from a small bakery. The bakery sells chocolate croissants for $0.70 each. The shop sells them for $2.40 to their customers. Unsold croissants can be returned to the bakery for $0.15 each. On the basis of the last few months, the shop expects a normal distributed demand for chocolate croissants. Expectation of demand is 14 croissants per day with a standard deviation of four per day.

$$c_o = 0.70 - 0.15 = 0.55\$$$

$$c_u = 2.40 - 0.70 = 1.70\$$$

$$CR = \frac{1.70}{1.70 + 0.55} = 0.75 \rightarrow F(0.7556) = 0.7 = z$$

$$S^* = 14 + 0.7 \cdot 4 = 17\ units$$

$$Z(S^*) = (0.55 + 1.70) f_{01}(0.7) 4 = 2.81\$/day$$

$$\Pi(S^*) = 1.70\ (14 - 2.81) = 2.99\$/day$$

Case study

This case study focuses on the impact of a global transportation concept on inventory management using the example of a shipping company.

The sea journey from two harbors takes 30 days on average; now, one departure of a vessel is offered each day instead of once a week. We assume a constant demand of 10 units a day, 98% service level, and a normally distributed lead-time from the origin to the destination harbor. The safety stock (ss) at destination can be calculated as:

$$ss = z\bar{d}\sigma_L$$

where z is the number of standard deviations, d is daily demand, and σ_L is standard deviation of lead time.

Assuming a service level of 98% and daily demand of 10 units, safety stock can be calculated as:

$$ss(weekly\ departure) = 2.055 \cdot 10 \cdot 2.0 = 41.1\ units$$

$$ss(daily\ departure) = 2.055 \cdot 10 \cdot 0.3 = 6.2\ units$$

The safety stock level falls significantly as a result of the company concept to only 6.2 units required in stock versus 41.1 units in a weekly service. This translates into lower inventory costs for direct customers or intermediaries, such as forwarders, at destination points and offers a measurable benefit to its customers.

Inventory Control Policies

Inventory control policy is a procedure that helps to define how much and when to order. The review may happen periodically (e.g., at the end of a month) or continuously (i.e., tracking each item and updating inventory levels each time an item is removed from inventory). Four parameters should be modelled:

- replenishment interval (t)
- order quantity (q);
- re-order point (s);
- target inventory level (S).

They can be fixed or changed (adjusted) in dynamics according to changes in demand and supply. Therefore, static and dynamic views on inventory control policies can be considered.

Order interval

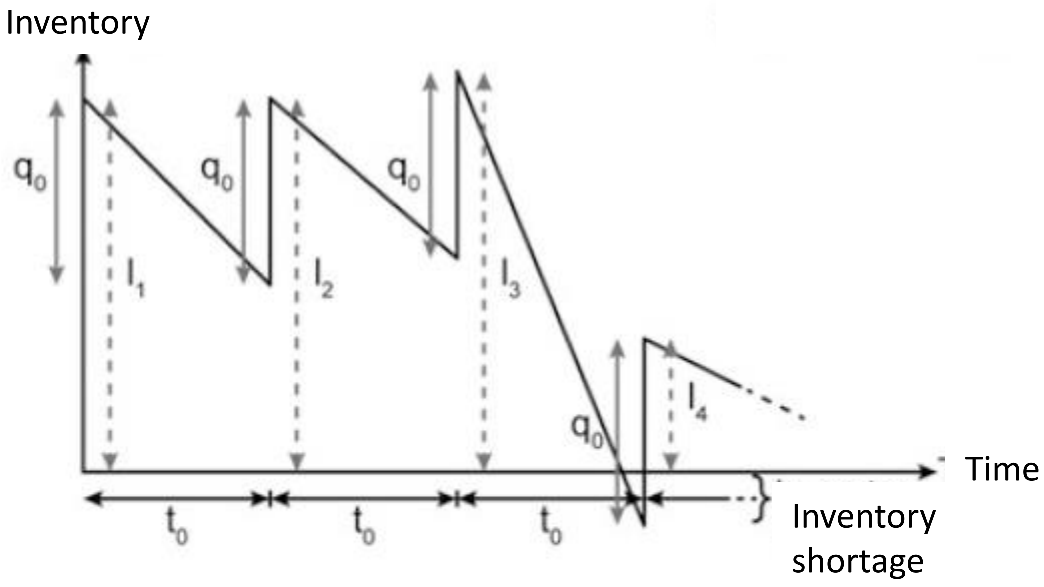
Order quantity	Fixed	Variable
Fixed	(t,q)-policy	(s,q)-policy
Variable	(t,s)-policy	(s,S)-policy

Order quantity and replenishment intervals may be both fixed and variable, therefore four basic inventory control policies can be classified.

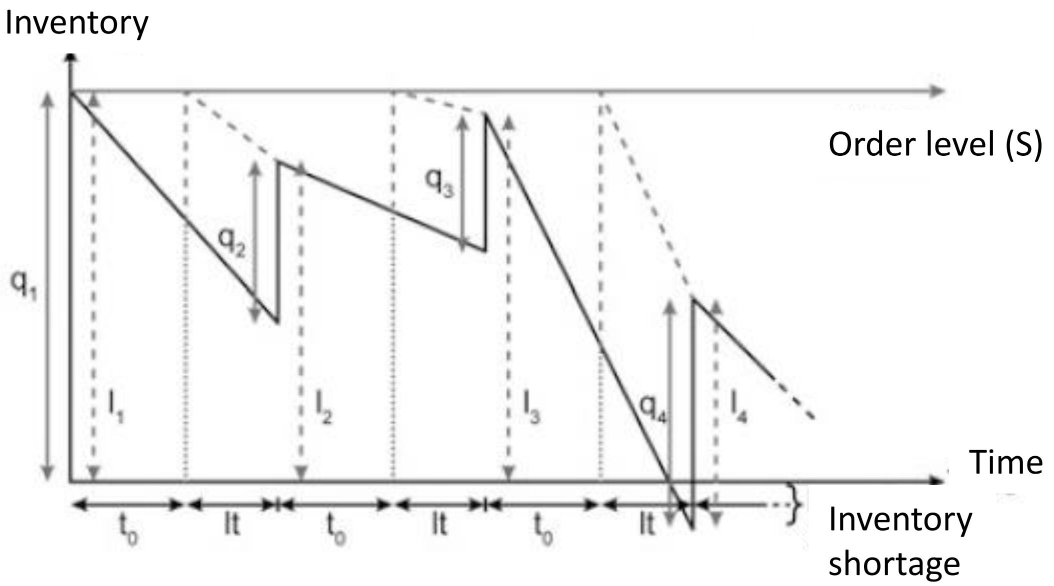
Fixed Parameters

When the replenishment interval, order quantity, ROP, and target inventory levels are fixed, the following policies can be classified.

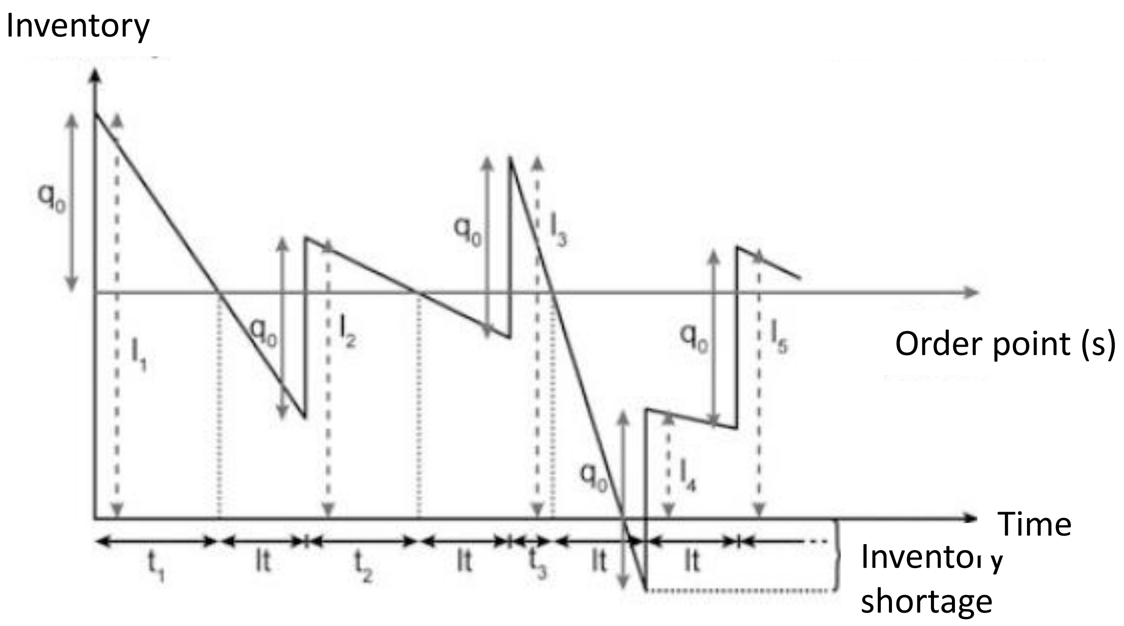
- Policy 1: t,q. A fixed amount (q) is ordered for a fixed period of time between two orders (t). It is recommended to implement this policy under constant demand, since it cannot be adjusted if uncertainty or fluctuation in demand exist.



- Policy 2: t,S. The order quantity (q) is variable, and q is placed at a fixed time (t). A certain amount of inventory needs to be ordered to reach the desired quantity S subject to lead time (lt). Order quantity is calculated as the target level S—stock on hand.

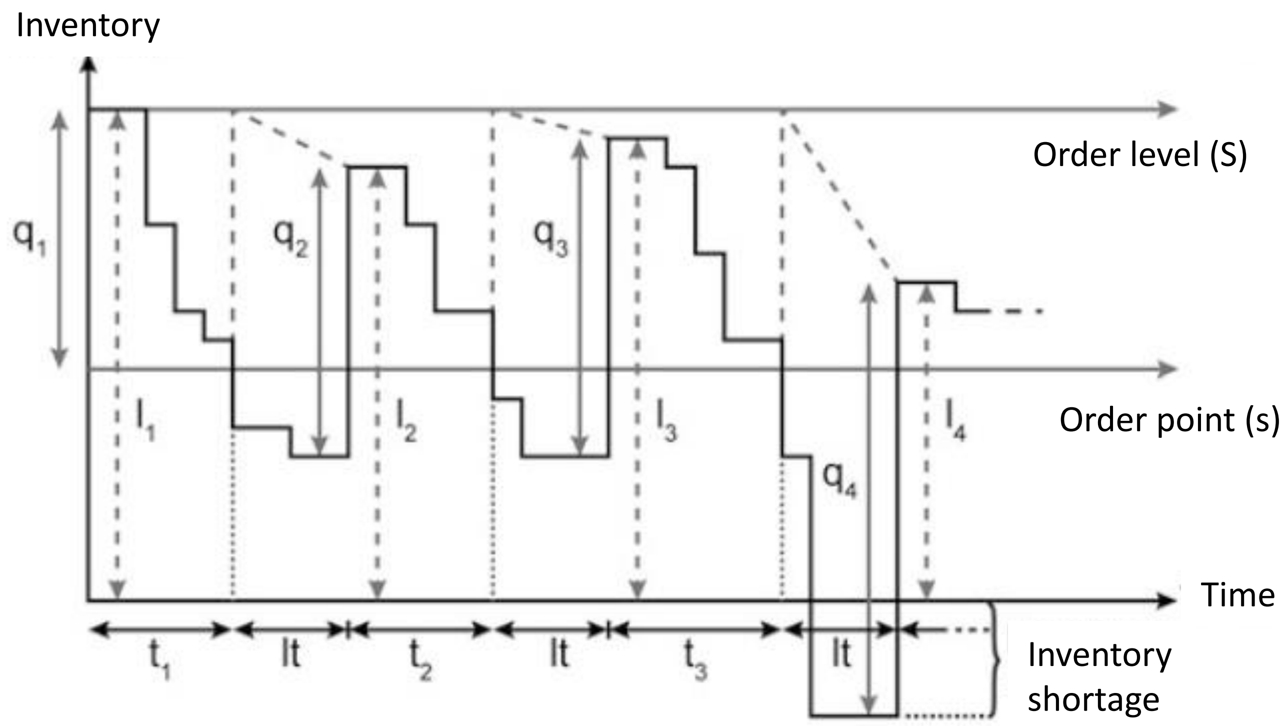


- Policy 3: s,q. This model operates when order quantity (q) is fixed and the interval (t) between orders can vary. In this case, the order point (s) is defined as ROP ($ROP = d(T + L) + z\sigma\sqrt{(T + L)}$). Every order arrives to replenish inventory after a lead time. The lead time is assumed to be known and constant. There is the possibility of shortage during an order cycle, that is, when the inventory level falls below zero (stock-out event). Every time inventory is extracted, what is left is compared with s. If the stock level is less than s, then an order at the rate of q is placed. Similar to the (t,q)-policy, in the (s,q)-policy, q also refers to the optimal order quantity.



- Policy 4: s,S. This strategy is used to define the drop of order quantity s after every inventory usage. In this case, inventory is to be refilled to raise the inventory position to the level S:

$$S = ROP + q$$



Example:

- demand per day (d): 100 units;
- standard daily deviation of demand (σ): 20 units;
- annual holding costs (h): $10 per unit;
- fixed ordering costs (f): $100 per order;
- order interval (T): 4 weeks;
- lead time (L): 2 weeks.

To determine parameters and annual holding costs for (s,S)- policy for 95% service level:

$$ss = z\sigma\sqrt{T+L} = 1.65 \cdot 20 \cdot \sqrt{4+2} = 81\, units$$

$$ROP = 100(4+2) + 81 = 681\, units$$

$$q^* = \sqrt{\frac{2 \cdot 36500 \cdot 100}{10}} = 855\, units \;\rightarrow S = 681 + 855 = 1536\, units$$

The policy is (681;1536). The average inventory position is (681+1536)/2=1108.

$$Costs = 1108 \cdot 10 = 11082\$$$

Dynamic view

When the replenishment interval, order quantity, ROP, and target inventory levels are not fixed, but change in dynamics subject to changes in demand, the following changes to the above-mentioned policies must be considered: demand, current and projected inventory, and in-transit quantities as well as planned deliveries.

Aggregating Inventory

In a number of cases, many markets are replenished from the same warehouse.

- A larger replenishment interval results in higher safety stock requirements
- A smaller replenishment interval means a lower level of inventory within the warehouse and thus a smaller warehouse is useful.
- A shorter replenishment interval probably results in higher labor cost because of greater material handling.

If two warehouses have to be merged in order to save warehousing fixed and operating costs. The right level of safety stock in the new larger warehouse can be determined as:

- If the two warehouses served the same market previously, safety stock should remain unchanged, according to the formula: $ss = z\sigma_{dLT}$
- If the two warehouses served different markets previously: $\sigma_{new} = \sqrt{\sigma_1^2 + \sigma_2^2}$

Example:
A firm selling milk has two warehouses. They serve different regions.
Demand for market A is 12,000 units with a standard deviation of 4600 units.
The market B has demand for 14,300 units with a standard deviation of 6200 units.
The purchase price for one unit of milk is 1.12 € and the retail price is 3.65 €. Milk units which are not sold can be sold on to a industry for 0.31 € per unit.
The company thinks about merging the two markets. Merging will involve extra transport costs of 22,500 €, but a reduction in fixed costs of 20,000 €.
To calculate optimal order volume, and expected costs and profit for every market, and decide if a merging the two markets is a profitable idea:

$$c_o = 1.12 - 0.31 = 0.81€$$
$$c_u = 3.65 - 1.12 = 2.53€$$
$$CR = \frac{2.53}{2.53 + 0.81} = 0.76 \rightarrow F(0.76) = 0.71 = z$$
$$f(0.71) = 0.31$$
$$S_A^* = 14300 + 0.71 \cdot 6200 = 18702$$
$$S_B^* = 12000 + 0.71 \cdot 4600 = 15266$$

Expected cost for A and B:

$$Z(S_A^*) = (0.81 + 2.53) \cdot 0.31 \cdot 6200 = 6420€$$
$$Z(S_B^*) = (0.81 + 2.53) \cdot 0.31 \cdot 4600 = 4763€$$

Expected profit for A, B and total:

$$\Pi(S_A^*) = 2.53 \cdot 14300 - 6420 = 29759€$$
$$\Pi(S_B^*) = 2.53 \cdot 12000 - 4763 = 25597€$$
$$\Pi(total) = 25597 + 29759 = 55356€$$

If A and B are merged:

$$\mu_{new} = \mu_A + \mu_B = 14300 + 12000 = 26300$$
$$\sigma_{new} = \sqrt{6200^2 + 4600^2} = 7720$$
$$S_{new}^* = 26300 + 0.71 \cdot 7720 = 31781$$
$$Z(S_{new}^*) = (0.81 + 2.53) \cdot 0.31 \cdot 7720 = 7993€$$

Comparison between the two solutions:

$$Z(S_A^*) + Z(S_B^*) - Z(S_{new}^*) = 6420 + 4763 - 7993 = 3190€$$
$$\Pi(S_{new}^*) = 2.53 \cdot 26300 - (7993 + 2500) = 56046€$$

Merging the two warehouses is a profitable solution.

Uncertainty is a system property characterizing the incompleteness of our knowledge about the system and the conditions of its development.

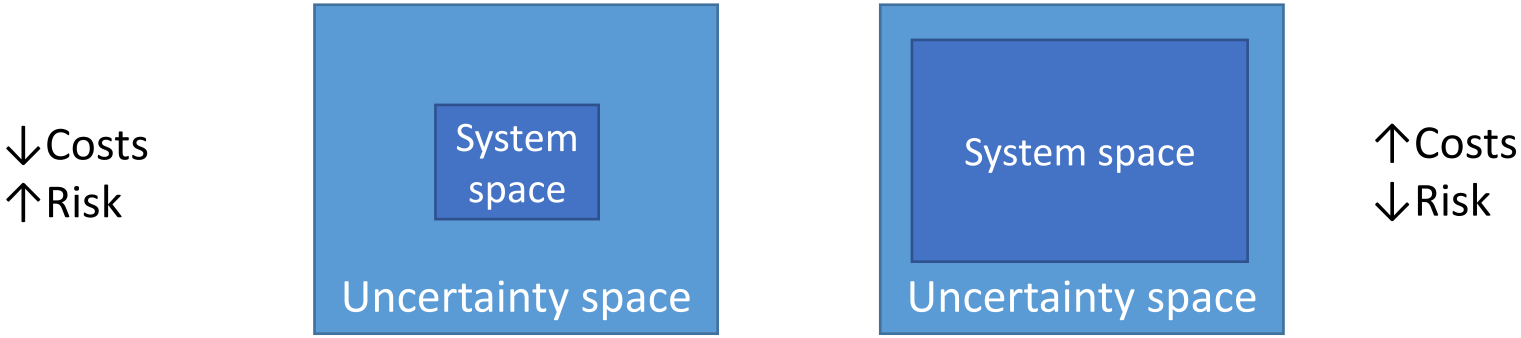
Risk is a measure of the set of possible (negative) outcomes from a single rational decision and their probabilistic values. The term “risk” is also replaced with “vulnerability”, which means “at risk”.

- Uncertainty is the general property of a system environment that exists independent of us for any system of a sensible complexity degree.
- Risk arises from uncertainty. Risks can be identified, analysed, controlled and regulated.
- A disturbance (perturbation impact) is the consequence of risks.
- It may cause a deviation (disruption) in the supply chain or not
- Operational deviations (or severe disruptions) are the result of perturbation influences. They may affect operations, processes, plans, goals or strategies. To adjust the supply chain in the case of deviations, adaptation measures need to be taken.

For the supply chain management domain, uncertainty factors and measures for their handling can be distinguished as:

Decision-making level	Uncertainty factors	Handling measures
Strategic	Multiple managementgoals Terrorism, piracy Financial and politicalcrises Natural disasters	**Multi-criteria analysis techniques supply chainsecurity management** **Liquid assets reserves Strategic material inventories** **Market diversification and outsourcing** **Product lines' flexibility and modularity**
Tactical and operational	**Weak coordination** **Stockless processes Weak control of cargosecurity** **Technological breaks** **Human errors**	**Safety stocks and time buffers** **Reserves of supply chain capacities supply chain coordination, monitoring, and event management**

The problem of a system under control and uncertainty is related to an area under control and an area under uncertainty:



By broadening the control area (right figure) and narrowing the uncertainty area or reverse (left figure), the system control can be adapted.

Example of risk classification:

- Disruptions (e.g. natural disasters, terrorism, war, etc.),
- Delays (e.g. inflexibility of supply source),
- Systems (e.g. information infrastructure breakdown),
- Forecast (e.g. inaccurate forecast, bullwhip effect, etc.),
- Intellectual property (e.g. vertical integration),
- Procurement (e.g. exchange rate risk),
- Receivables (e.g. number of customers),
- Inventory (e.g. inventory holding cost, demand and supply uncertainty, etc.),
- Capacity (e.g. cost of capacity).

- External risks which deal with threats from an external perspective of supply chain that can be caused by economical, sociopolitical or geographical reasons.
- Time risks referring to delays in SC processes
- Information risks, e.g., communication breakdown within the project team, information infrastructure complications, distorted information and information leaks
- Financial risks, e.g., inflation, interest rate level, currency fluctuations and stakeholder requests
- Supply risks, i.e., risks related to suppliers, e.g. supplier bankruptcy, price fluctuations, unstable quality and quantity of inputs
- Operational risks, caused by problems within the organisational boundaries of a firm, e.g. changes in design and technology, accidents and labour disputes
- Demand risks that refers to demand variability, high market competition, customer bankruptcy and customer fragmentation

Risk management is a methodological approach to managing uncertainty outcome.
People do not strive for a 100% guarantee of the result: they consciously tend to take risks.

- The risk factor is a global category that characterizes a system at the goal-orientation level (e.g. upsetting of the production plan, delivery breakdown, etc.).
- Risk sources consider certain events that may cause risk factors.
- The dangerous situation characterizes the state of a system when a probability of risk sources’ appearance and their direct influence on this system is high.
- The risk situation means a condition when the active influences of risk sources cause disturbances and deviations in system functioning.

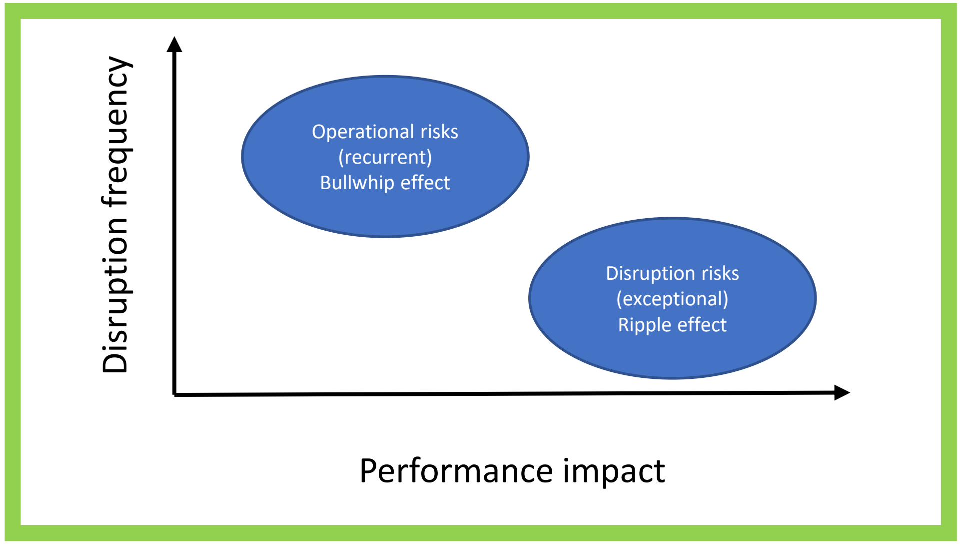
The problem of supply chain functioning in terms of risk consists of the following main phases:

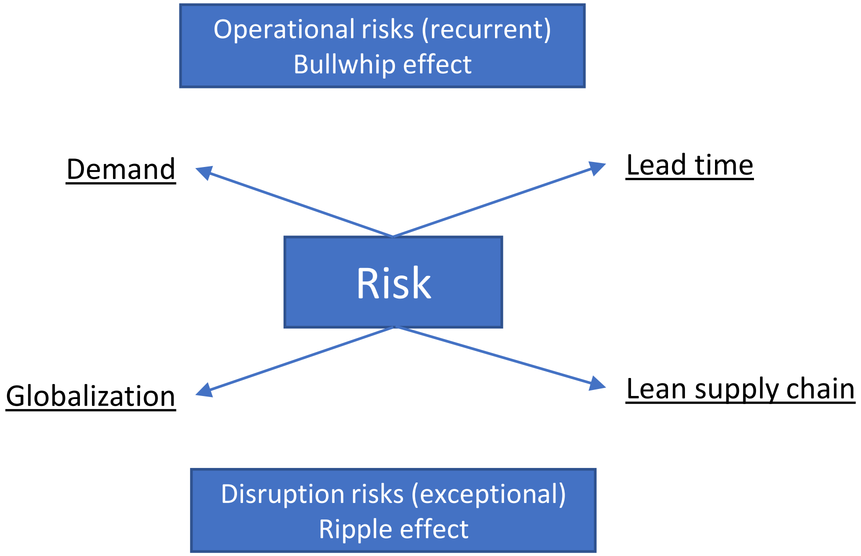
- risk factors’ identification
- risk sources and dangerous situations’ identification
- identification of interdependences between risk situation appearance and changes of system functioning parameters
- decision-making about compromise while supply chain configuration by aggravation of some goal criteria
- control decision development in order to compensate for possible disturbances in system functioning caused by risk situations development of a managed object monitoring system.

Uncertainties and risks in the supply chain can be furtherly classified as follows:

- random uncertainty (demand fluctuation risks)
- hazard uncertainty (risk of unusual events with high impact)
- deep uncertainty (severe disruption risks)

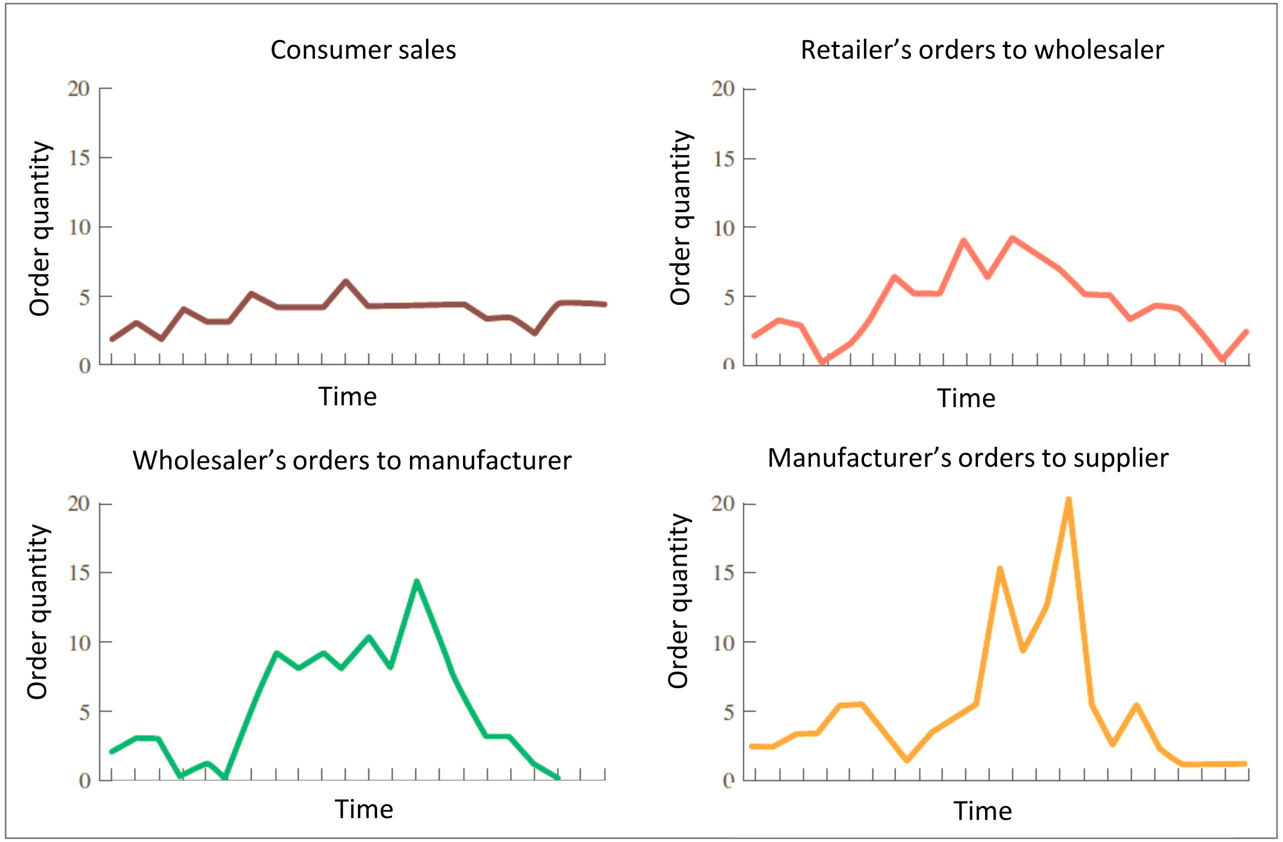
The different types of risks in the supply chain can be classified into demand, supply, process, and structure areas.





The Bullwhip effect considers weekly/daily demand and lead-time fluctuations as primary drivers of the changes in the supply chain which occur at the parametric level and can be eliminated in a short-term perspective.

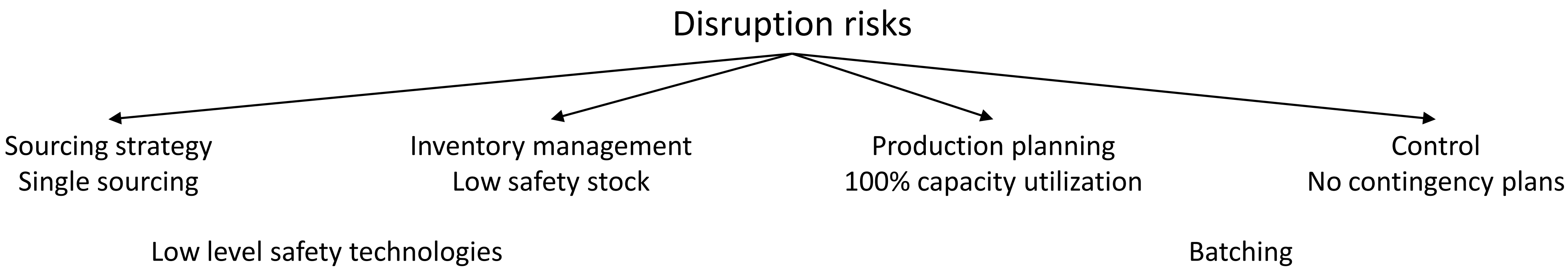
The Bullwhip effect considers weekly/daily demand and lead-time fluctuations as primary drivers of the changes in the supply chain which occur at the parametric level and can be eliminated in a short-term perspective. It refers to a situation where the variability in demand is magnified as moving from the customer to the producer in the supply chain.



The ripple effect in the supply chain:

- occurs if a disruption cannot be localized and cascades downstream impacting supply chain performance such as sales, stock return, service level, and costs.
- describes disruption propagation in the supply chain, impact of a disruption on supply chain performance and disruption-based scope of changes in supply chain structures and parameters.

The scope of the rippling and its impact on economic performance depends both on robustness reserves (e.g., redundancies like inventory or capacity buffers) and speed and scale of recovery measures.

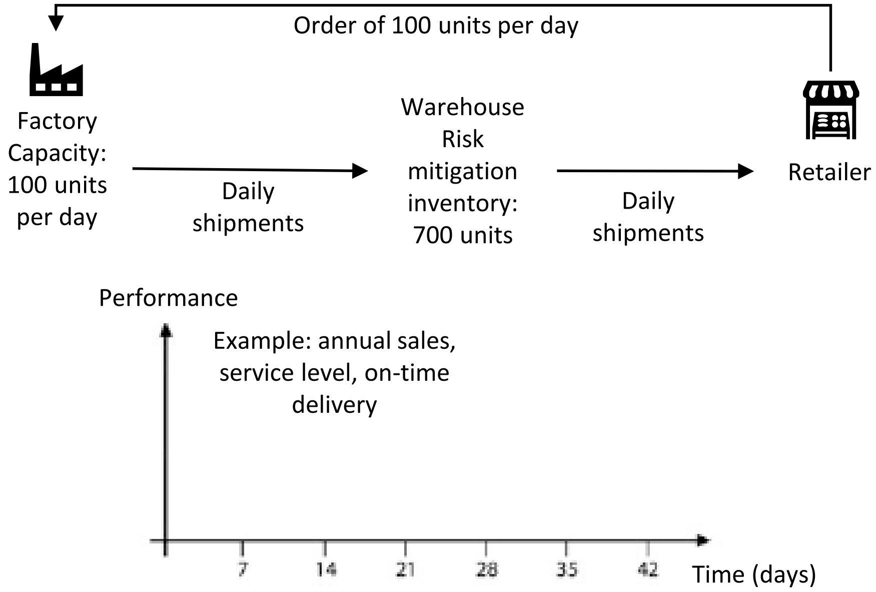


Supply chain resilience is the ability to maintain, execute and recover (adapt) planned execution along with achievement of the planned (or adapted, but yet still acceptable) performance is therefore the next objective property of the supply chain. Supply chains need to be planned to be stable, robust and resilient enough to:

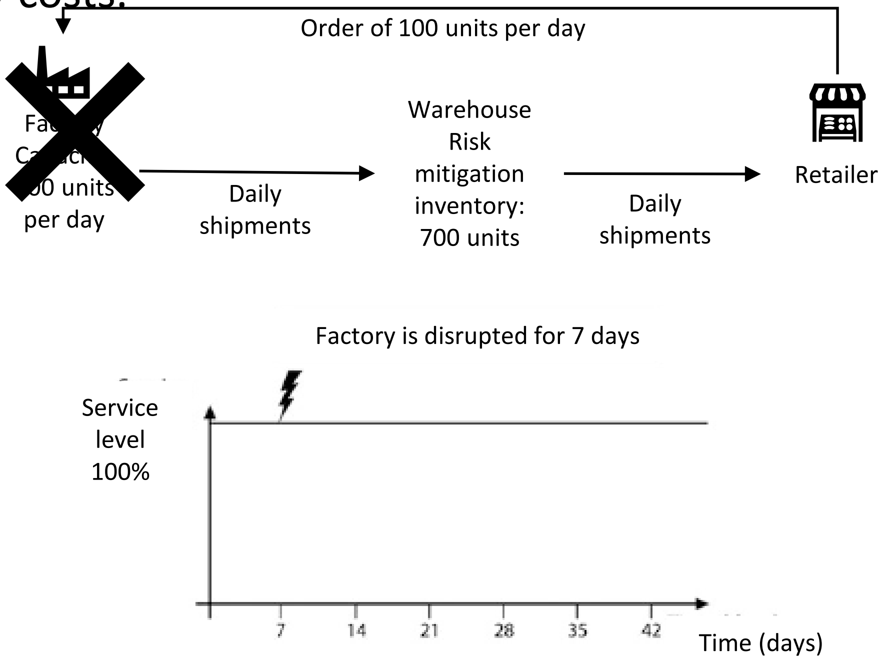
- maintain their basic properties and ensure execution
- be able to adapt their behavior in the case of disturbances in order to achieve planned performance using recovery actions.

- The robustness of supply chains is a complex characteristic of a non-failure operation, durability, recoverability, and the maintaining of supply chain processes and a supply chain as a whole. This is connected with the creation of a reserves system (the introduction of resource excessiveness) for the prevention of failures and deviations in supply chain processes.
- The flexibility of supply chains is a property concerning its ability to change itself quickly, structurally and functionally depending on the current execution state and reaching supply chain management goals by a change in supply chain structures and behaviour. This is connected with the creation of an adaptation system (with regard to operations and resources) for the prevention, improvement, or acquisition of new characteristics for the achievement of goals under the current environmental conditions varying in time.

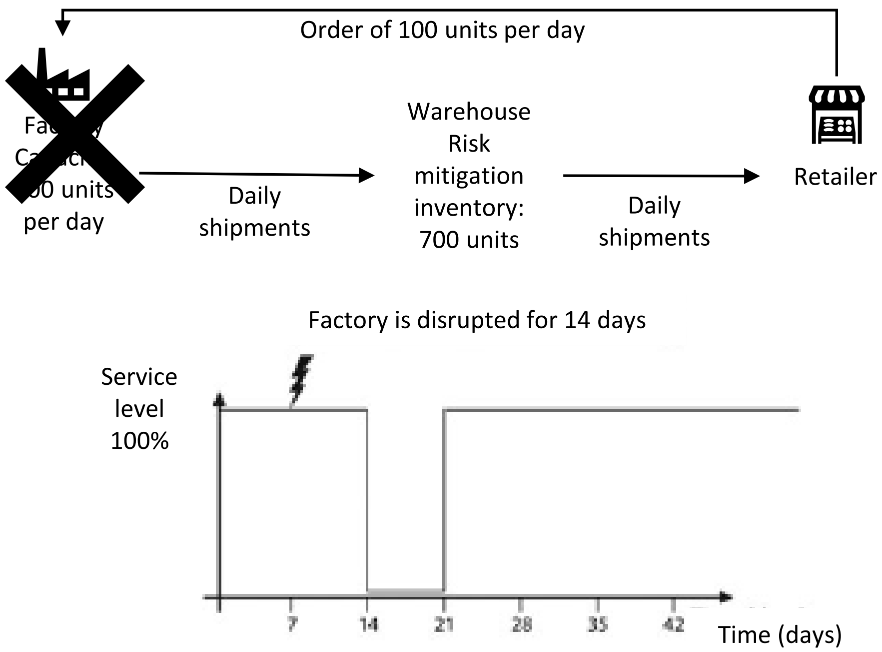
Increase in inventory, additional production capacities, and alternative transportation methods or back-up facilities would increase costs. At the same time, these so-called redundant elements would potentially lead to an increase in sales and service level. The robustness elements would also reduce risk of perturbations which may influence schedule execution. The resilient state of a supply chain requires a balanced robustness and flexibility which allows for achieving maximum service level with disruption risk considerations at acceptable redundancy costs.



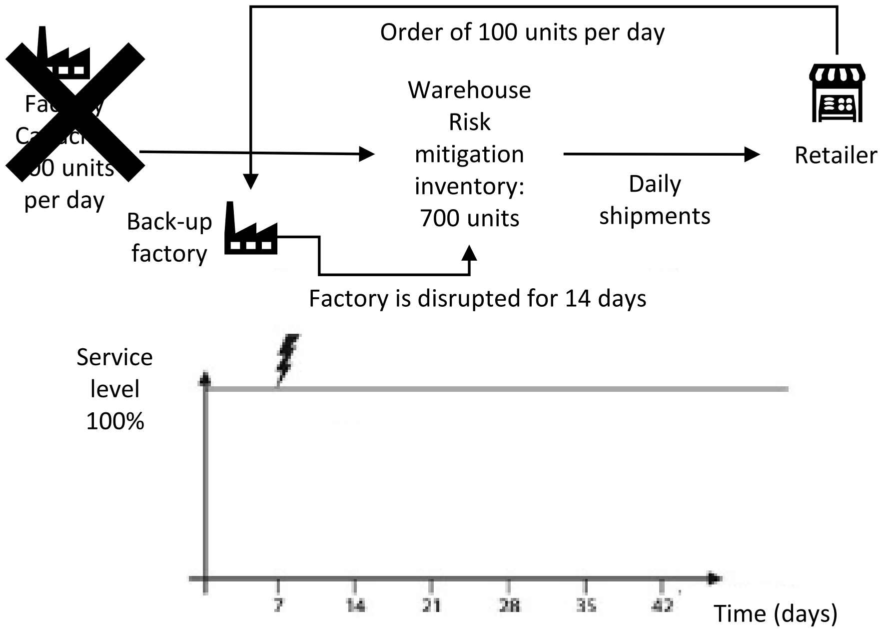
The factory orders 100 units of a product every day from factory that is aligned with the daily production capacity of 100 units. Daily shipments are assumed. The warehouse holds risk mitigation inventory of 700 units as a proactive resilience policy.



Factory capacity is disrupted for 7 days. Considering the risk mitigation inventory of 700 units and daily demand of 100 units, this disruption does not affect supply chain performance in terms of service level, i.e., the ratio of on-time delivered orders to all placed orders.

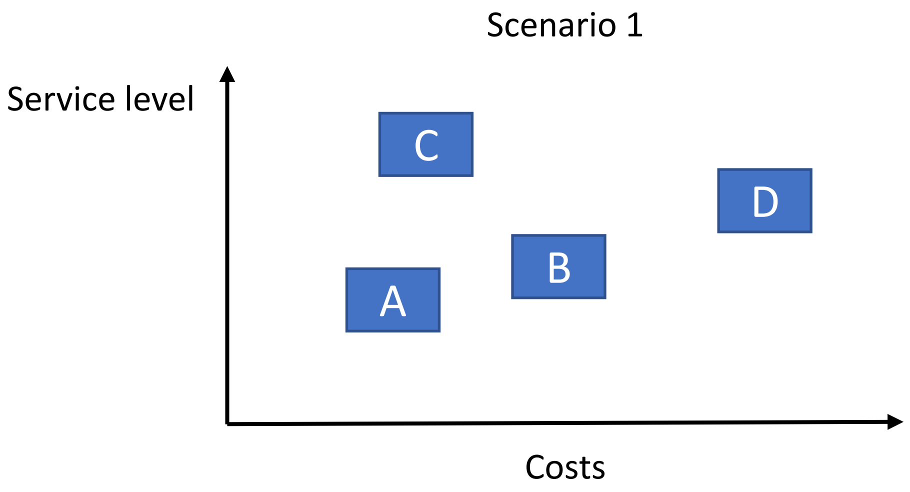


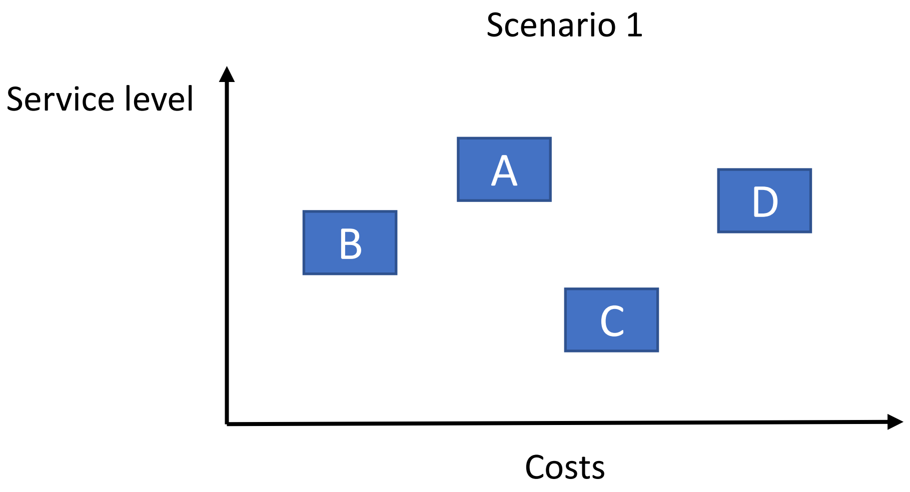
The factory stops producing for 14 days, service levels will be disrupted since risk mitigation inventory would help for 1 week only.



The usage of back-up factory mitigates the ripple effect and performance decrease. However, both risk mitigation inventory and a back-up capacity increase supply chain costs.

The balance of the elements of flexibility and robustness at proactive and reactive control loops, different constellations of service level, costs and stability should be analyzed.
The combinations of proactive mitigation strategies and reactive recovery policies may differ in costs and service level impacts for different disruption scenarios. The task of quantitative analysis methods is to decide on what proactive and reactive policies need to be selected.





Quantification of supply chain risk is based on KPIs (key performance indicators) that can be classified into operational and disruption risk KPIs with regards to bullwhip and ripple effects, respectively:

<u>Operational risk</u>

- The inventory variance ratio can be used to quantify the magnitude of bullwhip effect as the relation of inventory variance and demand variance ratio at a generic node:

$$InvVarR = \frac{\sigma^2_{inventory} / \mu_{inventory}}{\sigma^2_{demand} / \mu_{demand}}$$

where σ are variances of demand and inventory, and μ are the average values of demand and inventory.

- The order rate variance ratio is calculated as:

$$OrdVarR = \frac{\sigma^2_{order} / \mu_{order}}{\sigma^2_{demand} / \mu_{demand}}$$

$$\sigma^2_{demand} = \frac{\sum (x_{demand} - \mu_{demand})^2}{n-1} \quad \sigma^2_{order} = \frac{\sum (x_{order} - \mu_{order})^2}{n-1}$$

where x is the demand or orders in a period, μ is the average demand or orders, and n is the number of periods. If variance ratio measure is >1, then variance amplification is present and the risk of the bullwhip effect occurring arises. If the variance ratio measure is 1, then no bullwhip effect amplification is present. If the variance ratio measure is <1, then smoothing or dampening is occurring.

<u>Disruption risk</u>

- Ripple effect performance impact represents the relationship between the KPI planned (e.g., revenue or sales) in a disruption-free mode and the real KPI in the case of disruption:

$$PI = \frac{sales_{plan}}{sales_{disruption}}$$

Ideally, PI should be equal to 1. If the PI measure is >1, then disruption reduces supply chain performance. If PI measure is <1, then the supply chain planning was performed in a non-optimal manner.

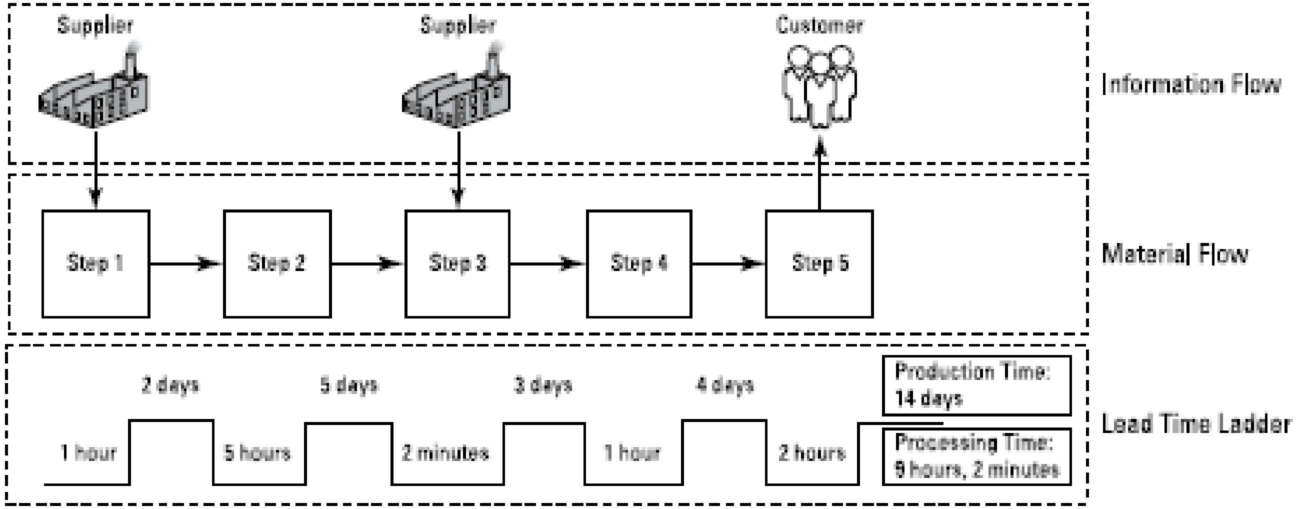
- Supplier reliability:

$$SR = \frac{quantity_{received}}{quantity_{ordered}}$$

- Time-To-Survive is the maximum time that the supply chain could operate and match supply with demand after a disruption
- Time-to-recovery is the time it would take for a particular node to be restored to full functionality after a disruption

In a supply chain network the goal of optimization is to deliver maximum value at the lowest cost. One way to achieve this goal is to change the nodes and the links.

Making changes in the links and nodes is called network optimization. One approach to network optimization is called value-stream mapping (VSM).



Supply chains are made up of people, processes, and technologies. All three components need to improve over time for a supply chain to remain competitive. People get better through education, training, and experience. Technology gets better through improvements in hardware and software. Processes get better through innovation and process improvement.

Three approaches to process improvement are particularly useful in supply chain management: Lean, Theory of Constraints, and Six Sigma.

Method	Focus
Lean	Reducing waste
Sigma	Reducing variability
Theory of constraints	Relieving bottlenecks

Lean

The idea behind Lean is that the least amount of time, effort, and resources is used by maintaining smooth and balanced flow in a supply chain. The best way to accomplish this goal is to have logical, disciplined processes and excellent communications. Everyone in the company needs to be working together to eliminate three things that cause inefficiency:

- Waste
- Unevenness or variability in operations
- Overburdening of people and equipment

Under the Lean approach, companies should continually drive eight kinds of waste out of their processes and supply chains:

- Transportation
- Inventory
- Motion
- Waiting
- Overproduction
- Overprocessing
- Defects
- Untapped skills and employee creativity

Six Sigma

Six Sigma is a process improvement method that's built on statistics. Consistent processes lead to a high quality level for products. The goal is to have a very small number of defects — that is, improved quality — as a result of decreased process variation.

The variation of a process is described in terms of the amount of deviation from an average value (σ). Any set of data about a process has some deviation, and the less deviation occurs, the more stable the process is. The statistical basis for Six Sigma is to reduce process variability so much that defects occur only at the sixth sigma (6σ).

Steps to apply Six Sigma as a process improvement methodology:

- The first step is to clearly define the process to improve and why improvement is needed (why the project is important and what resources is needed to complete it).
- The second step is measuring the process to improve. Because Six Sigma is a mathematical approach, data are to be collected to measure how the process is working and calculate the amount of variation.
- In the third step data are analyzed to identify variations in a process, show how those variations affect the quality of your products, and understand what things are causing the variability to look for ways to improve the process.
- In the fourth step changes to improve the process are made. These changes can happen all at once, or they can be introduced over time. Commonly, this phase includes some pilot studies to provide confirmation that the changes provide the expected benefits before implementing them throughout a process.
- The fifth step is establishing a system to ensure that the improvements you made become permanent. Control often involves performing ongoing measurements and reporting to show that the improvements remain in place and continue to provide consistency over time.

Theory of Constraints

The basic idea is that every process is limited by some kind of constraint. Theory of Constraints is about tuning an entire supply chain to run at the same pace as the slowest step in the process. The most restrictive step in a process is the one that constrains the entire system. The approach focuses improvement efforts on the constraints because that is where the greatest effect on the supply chain can be produced.

When constraints are detected, two approaches can be applied:

- Slow all the other steps so that they run at the same speed as the constraining step. This option prevents the buildup of inventory between the steps in the process.
- Improve the constraint so that the entire system moves faster. As constraint improvement continues, eventually, the constraint is no longer the slowest step in the process and stops being a constraint. Some other step becomes the constraint that's limiting the process, and the cycle starts again.

References:
• Watson, M., Lewis, S., Cacioppi, P., & Jayaraman, J. (2013). Supply chain network design: applying optimization and analytics to the global supply chain. Pearson Education.
• Chopra, S., & Meindl, P. (2016). Supply chain management: Strategy, planning, and operation. Boston, MA: Pearson.
• Jacobs, F. R., Chase, R. B., & Lummus, R. R. (2014). Operations and supply chain management. New York, NY: McGraw-Hill/Irwin.
• Ivanov, D., Tsipoulanidis, A., & Schönberger, J. (2019). Global supply chain and operations management: A decision-oriented introduction to the creation of value (Vol. 2). Cham, Switzerland: Springer International Publishing.