Dissertation on

"Big Data Analytics in Effective Inventory Management"

Thesis Submitted towards partial fulfilment

of the requirements for the degree of

Master in Production Management

Submitted by

DEBASHISH DAS

Examination Roll No: M4PRD19013

University Registration No: 140921 of 2017-2018

Under the Guidance of

Dr. Bijan Sarkar

Department of Production Engineering

Jadavpur University

Course affiliated to

Faculty of Engineering and Technology

JADAVPUR University

Kolkata-700032

India

2019

JADAVPUR UNIVERSITY

FACULTY OF ENGINEERING AND TECHNOLOGY

CERTIFICATE OF RECOMMENDATION

I HEREBY RECOMMEND THAT THE THESIS ENTITLED **"**BIG DATA ANALYTICS IN EFFECTIVE INVENTORY MANAGEMENT**"** CARRIED OUT UNDER MY SUPERVISION AND GUIDANCE, BY **MR. DEBASHISH DAS,** MAY BE ACCEPTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF **"MASTER OF PRODUCTION ENGINEERING".**

Countersigned

HEAD,

Production Engineering Dept.,

JADAVPURUNIVERSITY

KOLKATA-700032

Thesis Advisor

Professor,

Production Engineering Dept.,

JADAVPUR UNIVERSITY

KOLKATA- 700032

DEAN,

Faculty of Engineering and Technology,

JADAVPUR UNIVERSITY,

KOLKATA-700032

JADAVPUR UNIVERSITY

FACULTY OF ENGINEERING AND TECHNOLOGY

CERTIFICATE OF APPROVAL*

The forgoing thesis is hereby approved as a creditable study of an engineering subject carried out and presented in a manner of satisfactory to warrant its acceptance as a pre-requisite to the degree for which it has been submitted. It is understood that by this approval, the undersigned do not necessarily endorse or approve any statement made, opinion expressed and conclusion drawn therein but approve the thesis only for the purpose for which it has been submitted.

COMMITTEEON **(External Examiner)**

FINAL EXAMINATION

FOR EVALUATIONOF

THETHESIS **--**

(Internal Examiner)

*Only in case the recommendation is concurred in

ACKNOWLEDGEMENT

It is my greatest fortune to perform the thesis work in the Production Engineering Department, Jadavpur University.

It is great pleasure to express my gratitude and indebtedness to my esteemed guide Dr. Bijan Sarkar, Professor, Production Engineering Department, Jadavpur University. It is because of his continuous guidance, encouragement and valuable advice at every aspect and strata of the problem from the embryonic to the development stage that my thesis has been in the light of the day.

My special thanks to Head of Production Engineering Department for allowing me to carry out the research investigation with various facilities of the department. I would like to express my warmest gratitude to all the respected faculty members and non-teaching staff members of this department who directly or indirectly helped and encouraged me during the thesis work.

I would also like to thank to my elder brother Mr Sohag Das for his continuous support in completion of this work. I express my appreciation to my friends for their understanding, patience and active co-operation throughout my Masters course.

I feel pleased and privileged to fulfil my parents' ambition and I greatly indebted to them for bearing the inconvenience during my Masters course. Thank you, my beloved parents and my elder brother.

Everything in this nature is time bounded, so thanks to 'Almighty' for successful

Completion of the work in time.

Date:

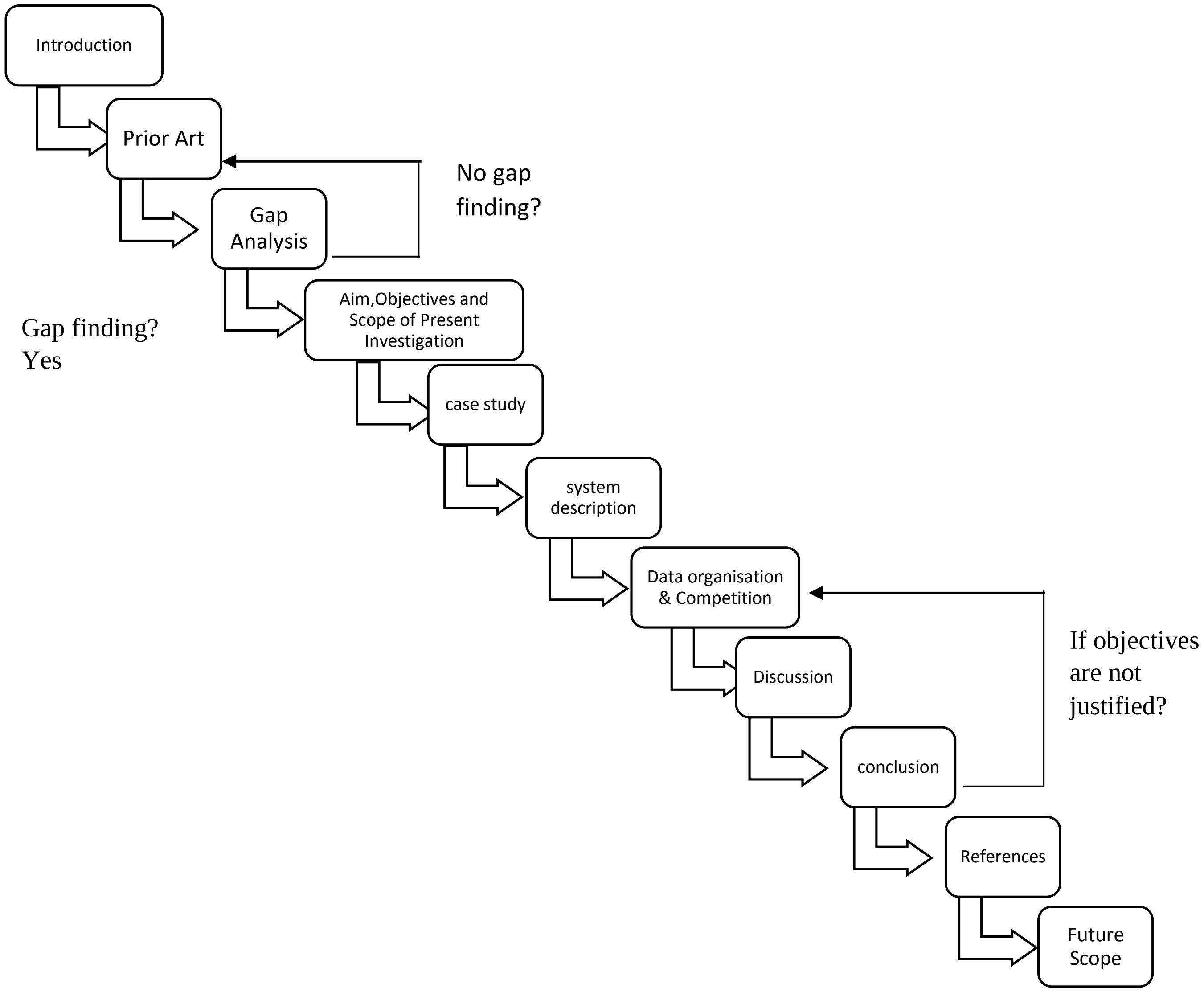
(DEBASHISH DAS)

Exam Roll No: M4PRD19013

Table of Contents

Page no

THESIS MAPPING:



LIST OF ABBREVIATIONS USED:

ABBREVIATED FORM	FULL FORM
ITR	Inventory Turnover ratio
EOQ	Economic Order Quantity
EPQ	Economic Production Quantity
RFID	Radio Frequency Identification
ROP	Reorder Point
Q	Order Quantity
P-models	Fixed Time-Period Models
GPS	Global Positioning System
SCM	Supply Chain Management
IoT	Internet of Things
SAS	Statistical Analysis System
MCDM	Multiple Criteria Decision Making
DM	Decision Maker
MOORA	Multi Objective Optimization on the basis of Ratio Analysis
AHP	Analytical Hierarchy Process
ANN	Artificial Neural Networks
DEA	Data Envelopment Analysis
TOPSIS	The Technique for Order of Preference by Similarity to Ideal Solution
SKU	Stock Keeping Unit
FIFO	First in First out
LIFO	Last in First out
SIRO	Service in Random Order
AAU	Average Annual Usage
CF	Critical Factor
AUC	Average Unit Cost
LT	Lead Time
IT	Information Technology

List of Figures

LIST OF TABLES

ABSTRACT

There are many things that determine the efficiency of a business. However, when that business involves the production of a tangible product, there is probably nothing more important to efficiency than the quality of inventory management.

In recent years there has been a revolution that is changing inventory management forever. That revolution is big data. Big data is the collection and analysis of a volume of digital information so vast that it couldn’t be stored on computer hardware until recently. It has transformed business analytics and been a game changer in many different industries.

Due to the increasing volume of transaction data and their correlated relations, it is often a non-trivial task to efficiently manage stocked goods, yet it is imperative to explore the underlying dependencies of the inventory items and give insights into implementing intelligent management systems. However, existing inventory management systems rely on statistical analysis of the historical inventory data, and have a limited capability of intelligent management. For example, they usually do not have the ability to forecast item demand and detect anomalous patterns of item inventory transactions. There is little work reported in implementing intelligent inventory management solutions to reveal hidden relations with integrated data-driven analysis. In this paper, I have tried to present the link between the inventory management and the big data analytics. A techniques like iMiner provides comprehensive support for conducting many inventory management tasks, such as forecasting inventory, detecting anomalous items, and analysing inventory aging. Our main focus is to find the new approach to eradicate the data ambiguity and predict the inventory management technique to usher a new dimension in the era of data diversity.

Chapter 1

1. INTRODUCTION

Inventory or stock (in common terms) is considered to be the central theme in managing materials. The inventory turnover ratio (ITR) is a barometer of performance of materials management function. In the generally understood term, inventory means a physical stock of goods kept in store to meet the anticipated demand. However, from materials management perspective, an apt definition of inventory is “a usable but idle resource having some economic value.”

It is necessary to have physical stock in the system to take care of the anticipated demand because no availability of materials when needed will lead to delays in production or projects or services delivered. However, keeping inventory is not free because there are opportunity costs of “carrying” or “holding” inventory in the organisation. Thus, the paradox is that we need inventory, but it is not desirable to have inventory. It is this paradoxical situation that makes inventory management a challenging problem area in materials management. It also makes a high inventory turnover ratio as a desirable performance indicator.

1.1 History of Inventory

We do not know precisely when inventory management arose. At its simplest, inventory management is about counting and keeping track of ‘things’. The earliest evidence archaeologists have found of humans counting ‘things’ are ancient tally sticks dating back approximately 50,000 years. Clay tokens found in Iran dating back over 4,000 years offer an interesting take on agricultural inventory; for example, to create a record representing two sheep, ancient ‘inventory managers’ would select two round clay tokens with + signs baked into them. Of course, using large numbers of tokens for very large flocks would be impractical, so different clay tokens were used to represent various numbers of different commodities.

Although we can draw a line between ancient counting systems and modern inventory management, that line is very long indeed. Ultimately, ancient inventory management was very basic and entirely manual. In many cases, the difficulty in counting items manually would mean that people would have to make inventory decisions based on a guess or a gut feeling.

TYPES OF INVENTORY

By the position in company`s production/operation process:

A. Raw materials
B. Works-in-process
C. Finished goods

By Estonian financial accounting rules:

A. Raw materials
B. Works-in-process
C. Finished goods
D. Goods purchased for resale
E. Advance payments to suppliers

Inventory Functions

- ✓ **Safety Stock**

An additional quantity of stock kept in inventory to protect against unexpected fluctuations in demands and/or supply. If demand is greater than forecast or supply is late, a stock shortage will occur. Safety stock is used to protect against these unpredictable events and prevent disruptions in manufacturing. Safety stock is also called buffer stock.

- ✓ ***Lot-size Inventory***

In order to take advantage of quantity price discounts, reduce shipping and setup costs, or address similar considerations, items are manufactured or purchased in quantities greater than needed immediately. Since it is more economical to produce or purchase less frequently and in larger quantity, inventory is established to cover needs in periods when items are not replenished.

- ✓ ***De-coupling Stock***

Inventory between facilities that process materials at different rates. De-coupling stock de-couples facilities to prevent the disparity in production rates at different facilities from interfering with any one facility’s production. This inventory increases the utilization of facilities.

- ✓ ***Pipeline Inventory***

Inventory to fill the transportation network and the distribution system including the flow through intermediate stocking points. This inventory exists because of the time needed to move goods from one location to another. Time factors involve order transmission, order processing, shipping, transportation, receiving, stocking, etc.

- ✓ ***Transportation Inventory***

Transportation inventory is part of pipeline inventory. It is inventory in transit between locations. The average amount of inventory in transit is:

$I = (A / 365) * D$

Where I is the average annual inventory in transit, A is annual usage, and D is transit time in days. The transit inventory does not depend upon the shipment size but on the transit time and the annual usage.

- ✓ ***Anticipation Inventory***

Additional inventory above basic pipeline inventory to cover projected trends of increasing sales, planned sales promotion programs, seasonal fluctuations, plant shut downs, and vacations. Anticipation inventory differs from safety stock in that it is a predictable amount.

- ✓ ***Hedge Inventory***

Inventory held to protect against future fluctuations due to a dramatic change in prices, strikes, war, unsettled government, etc. These events are rare, but such occurrences could severely damage a company's initiatives. Risk and consequences are usually high, and top management approval is often required. If the incident does not occur in the predicted time period, the hedge rolls over to the time period.

Safety, anticipation, and hedge inventories are compared in Table 1.

Table 1: Comparison of Safety, Anticipation, and Hedge Inventory

Inventory	Fluctuation	Time/Amount	Rolling Over
Safety	Unpredictable	Amount	No
Anticipation	Predictable	Time & Amount	No
Hedge	Semi-predictable	Time & Amount	Yes

Inventory Management

The word "Inventory" and "inventory management" have been defined in many ways in the literature. The following definitions are selected out of all as they are short and comprehensive:

"Inventory is a physical resource that a firm holds in stock with intent of selling it or transforming into a more valuable state."

"Inventory management is the branch of business management that covers the planning and control of the inventory. More specifically inventory management is a set of policies and controls that monitors levels of inventory and determines what levels should be maintained, when stock should be replenished and how large orders should be placed".

Needs of Inventory Management

There are several reasons that inspire companies to have stock. Five main reasons have been identified as motivation for holding inventory stocks, namely.

Economies of Sale:

The economies of scale in manufacturing, purchasing and Transportation Company can be realized by stock of inventory. A quantity discount is obtained if the business buys large amounts. Thus the transportation can move larger volumes and get economies of scale through better equipment utilization. If more materials is inventoried, manufacturing can have longer production runs allowing per unit fixed cost reduction.

Balance in supply in demand:

Inventory helps in maintaining a balance between supply and demand. A Christmas tree or newspaper manufacturer sees some demand year around but the demand increases by 60% or more in the Christmas Eve or any sensational news day. By manufacturing to stock, production can be kept level throughout the year. The idle plant capacity is reduced while maintaining a relatively stable workforce and also keeping the cost down. In the production of canned fruits, where the demand is relatively constant but the input materials are seasonal finished inventory helps meet demand when the materials are no longer available.

Specialization:

The subsidized firms can specialize with the help of inventory. Instead of manufacturing a variety of products, each plant can manufacture a product and then ship the finished products directly to customers or to a warehouse for storage. Thus by specializing, each plant can gain economies of scale through long production runs.

Protection from uncertainties:

In case of more demand and less supply of raw material stocks run out, the production line shuts down until more demand is delivered. Likewise a shortage of work in process means the product cannot be finished. Finally if a customer orders outstrips finished good supply, the resulting stock out could lead to loss of customers. Therefore the primary reason to hold inventory is to have protection from uncertainties.

Buffer Interface:

Inventory can buffer key interfaces, creating time and place utility. Key interfaces include –

1. Supplier and purchasing
2. Purchasing and distribution
3. Production and marketing
4. Marketing and distribution
5. Distribution and intermediary and
6. Intermediary and customer.

Having inventory at these interfaces helps ensure that demand is met and stock outs are minimized.

1.2 Application of inventory management

As inventory application grows in complexity, it's critical to maintain a strong understanding of its back-end functionality.

a) Maintain Applications Efficiently

The benefits of Inventory Management is the ability for an enterprise to interact smoothly and flexibly with their applications. It automates these mapping and application flow understanding tasks and carries them out with machine speed and efficiency, generating intuitive and highly useful data and program flowcharts that managers can reference for their own understanding. Enabled with this insight, managers save time that can better be used designing new features, developing more scalable applications environments, building stronger security protocols, and maintaining back-end functionality.

b) Reduce Downtime

A common cause for inventory application downtime is a misunderstanding or lack of understanding of inventory control. These are especially susceptible to significant downtime during fluctuating demand environment, volatile market demand, updates to foundational application elements, or just after release.

Inventory Management automatically maps this critical information quickly and insightfully, allowing managers with any amount of experience working with applications to address fatal runtime errors rapidly and effectively. This solution reduces planned downtime as well.

c) Address Security Vulnerabilities and Breaches

Removing vulnerabilities from your inventory applications requires a complete understanding of data sourcing, storage, referencing and passing in every instance in each of your supply chain network.

Vulnerabilities arise when calls for sensitive data are not vetted for caller authorization, when unnecessary volumes of data are passed or visible, when encryption keys are exposed, or when back doors to sensitive data exist.

d) Detect and Remove Obsolescence

It can consume computing resources wastefully and produce runtime errors. AIM obsolescence analysis is a novel way to automatically identify unused inventory, stock up in-process inventory and modules.

e) Make Informed Business Decisions

Sometimes, the biggest unknown in making business decisions around application changes is the impact, potential downtime, and costs of proposed changes. Empirical developer insight provides an approximate picture, at best, of the impact of a change proposal on inventory application.

Inventory Management enables your business to perform a thorough impact analysis on Inventory application architecture to get a complete picture of the impact of a change on Inventory application. Not only does it generate an analysis quickly and on-demand, it provides pinpoint accuracy as to all of the related pieces of Inventory application that will be affected by a change. A good team of developers will build and maintain a working application flow map for reference to maintain your application effectively. Inventory Management automates this task and provides insight in a format that is useful for managers freeing your workers up to keep your Inventory Management applications up, running, and evolving.

1.3 Various inventory management techniques

The first mathematical inventory model is generally referred to as the Economic Order Quantity (EOQ) model which was developed by Harris in 1915. Several successful attempts have been made to use the various extensions of EOQ in practice. Further research works showed that the EOQ model appears to be quite insensitive to errors in the specification of the appropriate cost parameters and the estimation of demand, the importance of the EOQ model is not only from the historical point of view but also because many other models designed to cope with different situations have been based on this very model.

The following models are discussed over here to get an overview of the general principles of inventory management.

Deterministic Models:

- The Economic Order Quantity(EOQ) model
- The Economic Production Lot Size Model
- An inventory model with planned shortages

Probabilistic model:

- Single-period inventory model
- A continuous fixed order Quantity Model
- A Fixed time period Model

1.4 Inventory Costs

There are four types of costs that must be considered in setting inventory levels.

Purchasing cost:

This is cost of purchasing or procuring inventory. Inventory is a liability from the point view of expenditure by means of locking working capital. To have inventory the firm mush have sufficient working capital. The unit cost of items purchased many fluctuate over the entire planning horizon. Hence, average unit price is considered for computing the purchasing cost of inventories.

Holding costs:

Holding cost or carrying costs are costs such as storage, handling, insurance, taxes, obsolescence, theft and interest on funds financing the goods. These charges increase as inventory levels rises. In order to minimize carrying costs management makes frequent orders of small quantities. Holding costs are commonly assessed as a percentage of unit value, that is l5 percent, 20 percent, rather than attempting to derive a monetary value for each of these costs individually. This practice is a reflection of the difficulty inherent in deriving a specific per unit costs for example theft or obsolescence.

Ordering costs:

Ordering costs are those costs associated with placing an order, including expenses related to personnel in a purchasing department, communications and the handling of the related paperwork. Lowering these costs would be accomplished by placing a small number of orders, each for a large quantity. Unlike carrying costs, ordering costs are generally expressed as a monetary value per order.

Stockout costs:

It included sales that are lost, both short and long term. These charges are probably the most difficult to compute, but arguably the most important because they represent the costs incurred by customers (internal or external) when inventory policies falter. Failure to understand these costs can lead management to maintain higher (or lower) inventory levels than customer requirement may justify.

1.5 Basic Inventory Models

The first mathematical inventory model is generally referred to as the Economic Order Quantity (EOQ) model which was developed by Harris in 1915. There are several full length books attempted to explain how various extensions of EOQ can be used in practice. Further research works showed that the EOQ model appears to be quite insensitive to errors in the specification of the appropriate cost parameters and the estimation of demand. The importance of the EOQ model is not only from the historical point of view but also because many other models designed to cope with different situations have been based on this very model.

The following models are discussed over here to get an overview of the general principles of inventory management.

1.5.1 Deterministic Models:

1.5.1.1 The Economic Order Quantity Model:

Overview:

The Economic order quantity (EOQ) applies only when demand for a product is constant over the year and each new order is delivered in full when inventory reaches zero. There is a fixed cost for each order placed, regardless of the number of units ordered. There is also a cost for each unit held in storage, commonly known as holding cost, sometimes expressed as a percentage of the purchase cost of the item.

The optimal number of units to order needs to be determined so that the total cost associated with the purchase, delivery and storage of the product can be minimized.

The required parameters to the solution are the total demand for the year, the purchase cost for each item, the fixed cost to place the order and the storage cost for each item per year. Note that the number of times an order is placed will also affect the total cost, though this number can be determined from the other parameters.

Assumptions:

- Demand for items from inventory is continuous and at a constant rate.
- Orders to replenish inventory are made at regular intervals
- Ordering cost is fixed (independent of quantity produced)
- The lead time is fixed.
- The purchase price of the item is constant, i.e. no discount is available.
- The replenishment is made instantaneously

Total minimum cost model:

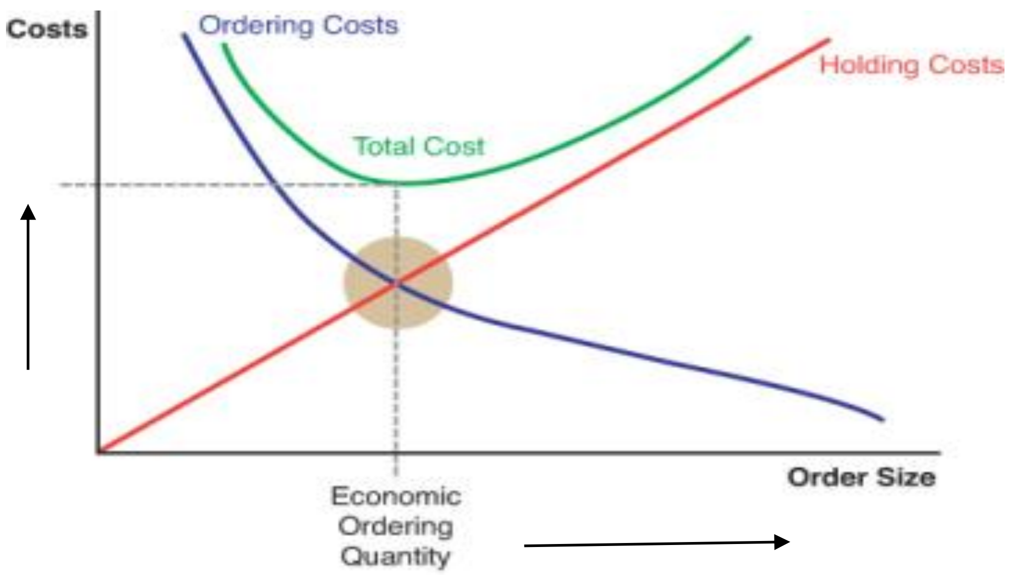


Fig 1: Annual Holding, Ordering, Tool cost & EOQ

Total cost= Purchase cost or production cost +Ordering cost +Holding cost

Where: Purchase cost: this is the variable cost of goods: purchase unit price ×annual demand quantity & this is P × D.

ORDERING COST: this is the cost of placing orders: each order has a fixed cost of K, and it needs to order D/Q times per year. This is K×D/Q

Holding cost: the average quantity in stock (between fully replenished and empty) is Q/2, so this cost is h × Q/2.

TC = PD+DK/Q +hQ/2

To determine the minimum point of the total cost curve, calculate the derivative of the total cost with respect to Q (assume all other variables are constant) and set it equal to 0:

$$0 = -\frac{DK}{Q^2} + \frac{h}{2}$$

Solving for Q gives Q* (the optimal order quantity):

Q^{*2}= 2DK/h

Formula:

D-Annual demand, Q- order quantity, k-cost of placing order, h- Annual per unit holding cost, ordering cost= KD/Q, Holding cost= HQ/2,

Total cost= KD/Q + hQ/2

$$Q *= \sqrt{2DK/h}$$

Graphical Representation:

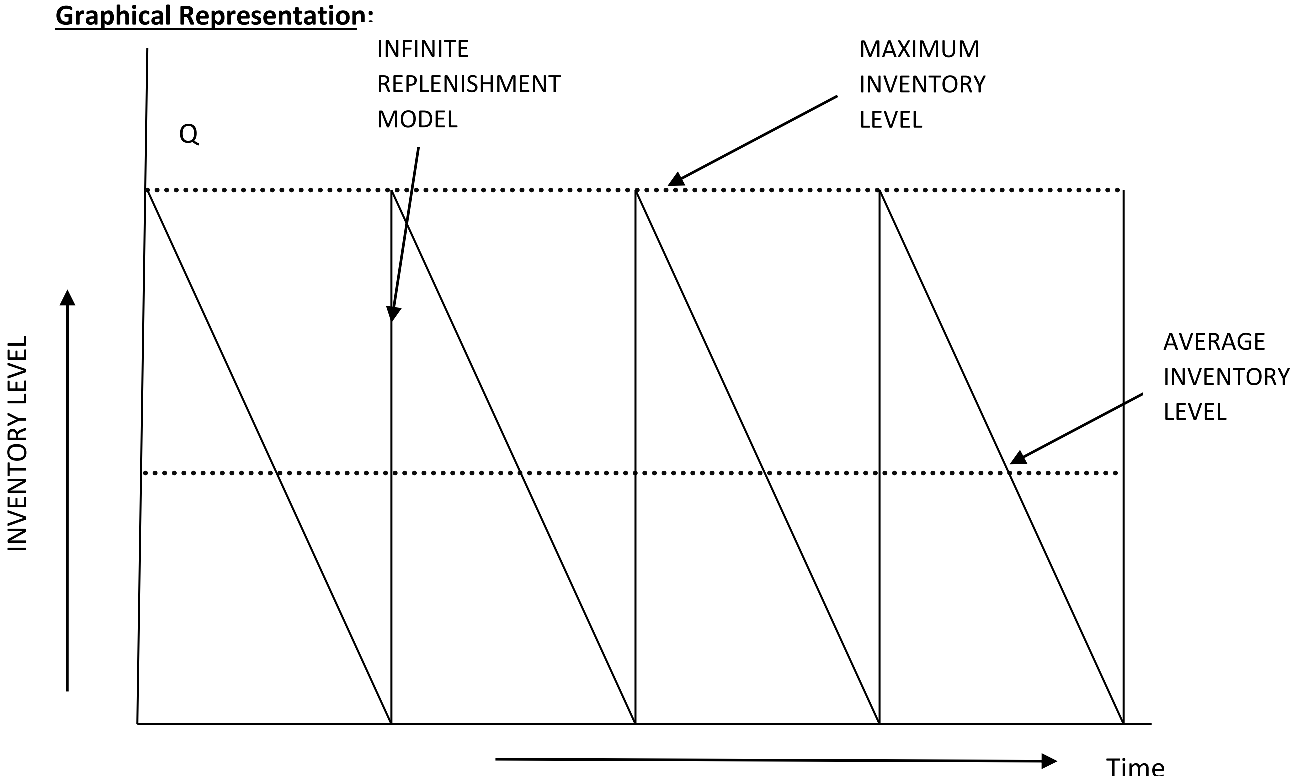


Fig. 2: Inventory pattern for EOQ Inventory Model

1.5.1.2 The Economic Production Lot Size Model

Overview:

Economic Production Quantity (EPQ) or the Economic Production Lot Size Model only applies where the demand for a product is constant over the year and that each new order is delivered/produced incrementally when the inventory reaches zero. There is a fixed cost charged for each order placed, regardless of the number of units ordered. There is also a holding or storage cost for each unit held in storage (sometimes expressed as a percentage of the purchase cost of the item).

Here also, the optimal number of units of the product to order is determined so that the total cost associated with the purchase, delivery and storage of the product is the minimum.

The required parameters to the solution are the total demand for the year, the purchase cost for each item, the fixed cost to place the order and the storage cost for each item per year. The number of times an order is placed will also affect the total cost, however, this number can be determined from the other parameters.

Assumption:

- Demand for items from inventory is continuous and at a constant rate
- Production runs to replenish inventory are made at regular intervals
- During a production run, the production of items is continuous and at a constant rate
- Production set-up/ordering cost is fixed (independent of quantity produced)
- The lead time is fixed
- The purchase price of the item is constant, i.e. no discount is available
- The replenishment is made incrementally

Variables:

K = ordering/setup cost, D = demand rate, h = holding cost, T = cycle length,

P = production rate, $x = \frac{D}{P}$, Q = order quantity

$\text{Holding Cost per Year} = \frac{Q}{2} \times h(1-x)$,

Where $\frac{Q}{2}$ is the average inventory level, and h (1 — x) is the average holding cost. Therefore multiplying these two results in the holding cost per year.

Ordering Cost per Year = $\frac{D}{Q}$ X K

Where, $\frac{D}{Q}$ is the orders placed in a year, multiplied by K results in the ordering cost per year.

EPQ Formula

$$Q^* = \sqrt{\frac{2KD}{h(1-x)}}$$

<u>Graphical representation</u>

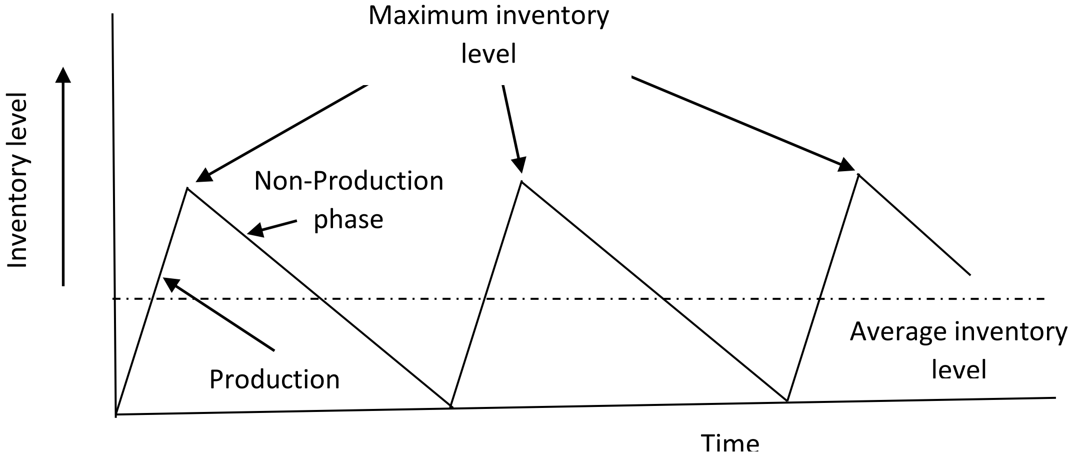


Fig. 3: Inventory Pattern for Production Lot Size Inventory Model

1.5.1.3 An Inventory Model with Planned Shortages

Overview

One of the assumptions of our basic EOQ model is that shortages and back ordering are not allowed. The third model variation that will be described, the EOQ model with shortages, relaxes this assumption. However, it will be assumed that all demand not met because of inventory shortage can be back ordered and delivered to the customer later. Thus, all demand is eventually met. The EOQ model with shortages is illustrated in Fig.

Assumptions:

- All assumptions made EOQ model with the following exception:
- Shortages are allowed as backorder assuming no lost sales.

Formula:

Bypassing the lengthy derivation of the individual cost components of the EOQ model with shortages, which requires the application of plane geometry to the graph in below Figure. The individual cost functions 'are provided as follows, where S equals the shortage level and C_s equals the annual per-unit cost of shortages, $C_o$equals the unit ordering cost, C_c equals the annual per-unit holding/carrying cost:

$$\text{Total shortage costs} = C_s \frac{S^2}{2Q}$$

$$\text{Total carring costs} = C_c \frac{(Q-S)^2}{2Q}$$

$$\text{Total ordering costs} = C_o \frac{D}{Q}$$

Combining these individual cost components results in the total inventory cost formula:

$$TC = C_s \frac{S^2}{2Q} + C_c \frac{(Q-S)^2}{2Q} + C_o \frac{D}{Q}$$

The only way to determine the optimal order size and the optimal shortage level, S, is to differentiate the total cost function with respect to Q and S , set the two resulting equations equal to zero, and solve them simultaneously . Doing so results in the following formulas for then optimal order quantity and shortage level;

$$Q_{opt} = \sqrt{\frac{2C_0 D}{C_c} \frac{(C_s + C_c)}{C_s}}$$

$$S_{opt} = Q_{opt}(\frac{C_c}{C_s + C_c})$$

Graphical Representation

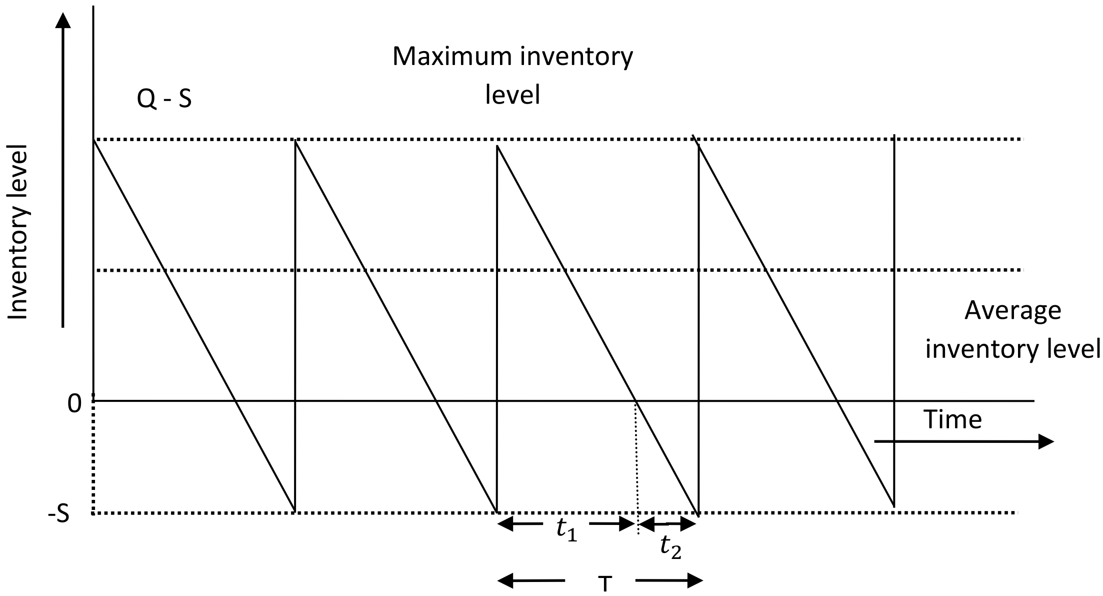


Fig.4: Inventory Pattern for EOQ Inventory Model with Back Orders

1.5.2 Probabilistic Models:
1.5.2.1 Single-period inventory model

Overview

It is necessary to clarify the term single period. This term refers to the situation where the inventory is perishable and demand for that particular inventory exists only for the Period which it is ordered (or) procured. Newspaper selling is such an example. The newspaper ordered for today will not be sold at the same price tomorrow. Fashion selling is another example Spring-summer designs will not sell during the autumn-winter season.

Derivation

Increment analysis is used to determine the optimal order quantity for a single-period inventory model with probabilistic demand. The increment analysis addresses the how-much-to order question by comparing the cost or loss of ordering one additional unit with the cost or loss of not ordering one additional unit. Notation used in this model is listed below.

Co: Cost per unit of overestimating demand; represents the loss of ordering one additional unit that may not sell.

Cu: Cost per unit of underestimating demand; represents the loss of not ordering one additional unit for which demand existed otherwise.

Let the probability of the demand of inventory being more than a certain level y is P (D > y), and the probability of the demand of inventory being less than or equal to this level y is P (D <= y). Then, the expected loss (EL) is given by either of the two conditions below.

Overestimations: EL(y + 1) = C0 * P (D <= Y)

Underestimation: EL(y) = Cu * P (D > Y)

Following which the optimal order quantity (y *) can be found as follows: EL(y * + 1) =EL(y*)

Formula

Co * P (D <= y*) = Cu * P (D > y*); it is known that P (D > y*) = 1 - P (D <= y*).

Substituting above two equations, it becomes Co * P (D <= Y *) = C * [1 - P (D <= y *)]

Solving for P (D <= y *), it is finally obtained that P (D <= y*) = $\frac{C_u}{C_c+C_u}$

The above expression provides the general condition for the optimal order quantity Y * in the single-period inventory model. The determination of y * depends on the probability distribution.

1.5.2.2 A Continuous Fixed Order Quantity Model

Overview

In earlier periods, non-continuous, or periodic inventory systems were more prevalent. This has been facilitated by bar coding and lately radio frequency identification (RFID) labelling which allows computer systems to quickly read and process inventory information as part of transaction processing.

The reorder point for replenishment of stock occurs when the level of inventory drops down to zero. In view of instantaneous replenishment of stock the level of inventory jumps to the original level from zero level.

Reorder Point, ROP = Normal consumption during lead-time + Safety stock

Several factors determine how much delivery time stock and safety stock should be held

In summary, the efficiency of a replenishment system affects how much delivery time is needed.

Another method of calculating reorder level involves the calculation of usage rate per day, lead time which is the amount of time between placing an order and receiving the goods and the safety stock level expressed in terms of several days' sales.

Reorder level = Average daily usage rate x lead-time in days = D * L + Z * α (L), where D is average demand during lead time, L = Lead Time, Z = standard normal variant according to the desired service level, α (L) = standard deviation of demand during lead time.

Features:

- Order Quantity (Q) - Fixed, the same amount is ordered each time, less error-prone and paper work is reduced.
- When to place an order- When inventory position drops to the reorder point (ROP).
- Recordkeeping*- Each time a withdrawal or addition is made. With the advent of modern computer systems and inventory records policy, this effort is much reduced
- Size of inventory- Less than Fixed-Time Period Model because it provides safety stock only for the lead time.

- Time to maintain - Minimal because the point-of-sale scanners update inventory each time a sale is made.
- Type of items- Higher priced, critical, or important items (e.g. diamonds, computer chips etc.) because average inventory.

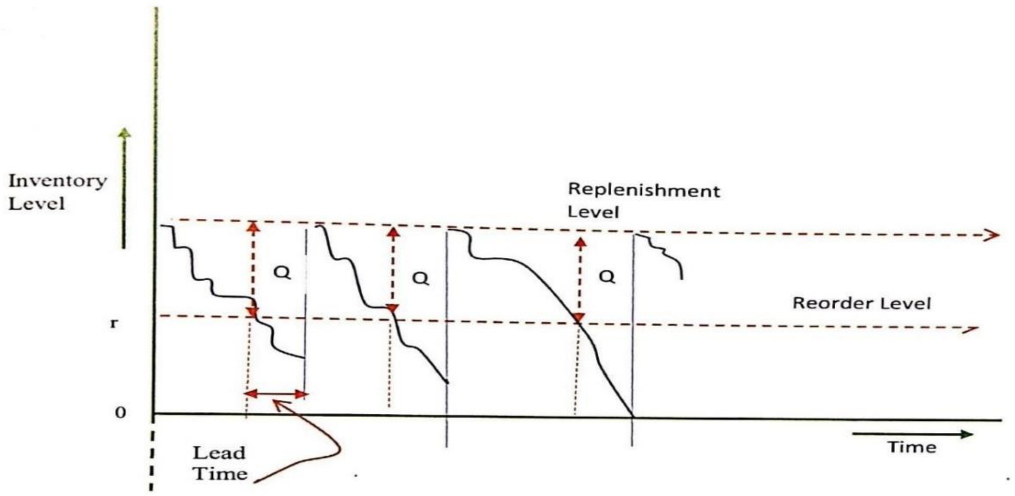


Fig.5: Inventory Pattern for Continuous Review Model with Probabilistic Demand

1.5.2.3 A Fixed Time Period Model

Overview

In fixed time-period models (also known as P-models), orders are placed at fixed periods, irrespective of the demand or usage pattern. An order quantity to replenish available inventory to a maximum level is placed. Because of the uncertain demand pattern, a safety stock is usually maintained; safety stock is the minimum stock levels maintained, which is not accounted for in evaluating the order quantity. The model is illustrated in below Fig. The order quantity in this model is dependent on demands, safety stock and current inventory.

Q=D * (T + L) + Z * α (T + L) - I where, D is the average demand T is the periodicity of review, L is the lead time, Z is the number of standard deviations for a specified service probability, α (T + L) is the standard deviation of demand over review and lead time, I is the current inventory (including those being processed in order), $\alpha\ (\mathrm{T}\ +\ \mathrm{L}) = \sqrt{\sum_{i=1}^{T+L} \alpha_{(D_i)^2}}$

Now,$\alpha_{(D_i)}$ can be assumed to be constant, as demands are considered independent over any period. Therefore, SS = Z * α (T + L) = Z × $\sqrt{(\mathrm{T}\ +\ \mathrm{L})\ *\ \alpha_L^2}$

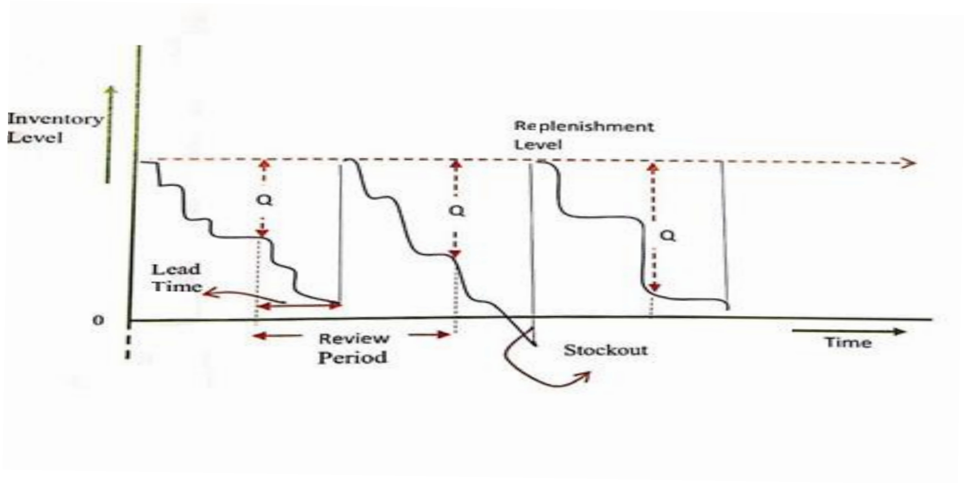


Fig.6: Inventory Pattern for Periodic Review Model with Probabilistic Demand

The safety stock to be maintained is SS = Z X$\sqrt{(T + L) * \alpha_L^2}$

1.6 Inventory Classification

1.6.1 Selective Control Policies

Inventory in a company Consists of thousands of different items in stock. The control of all these items creates a serious problem to the management if the same amount of control is exercised on each of these items.

Therefore, in order to execute proper control, it is necessary to take selective approach and find the attention required for each item according to its importance. This is essential for achieving maximum benefits with minimum efforts and costs.

Depending upon the alternatives and purposes, different analysis have been developed to help in bringing practical solution to the problem of inventory control. The commonly used systems can be classified as:

The ABC Classification

It is usually uneconomical to apply detailed inventory control analysis to all items carried m an inventory. Frequently, a small percentage of inventory items accounts for most of the total inventory value. It is usually economical to purchase a large supply of low cost items and maintain little control over them. Conversely, small quantities of expensive items are purchased, and tight Control is exercised over them.
This method is based on the annual value consumption figures. This analysis concentrates all stocked items into three uneven classes. The reason for this type of grouping is that most of

the organizations, it is seen that a large proportion of the usage or sales value of stocked items is concentrated in only relatively few of those items. There are other items the number and consumption value lie between A and C groups. These are B items. That is, A= high value items, B = medium value items, C = low value items. The method is as follows: - The inventory value for each item is obtained by multiplying the annual demand by the unit cost. The entire inventory is lifted in descending order from the largest value to the smallest. The items are then designated by the ABC classification system. The cumulative consumption values are then computed in the same order.

Table 2: ABC classification empirical rule (Pareto principle)

Category	% of items (approx.)	% of Usage Value (approx.)
High Value Items	20%	80%
Medium Value Items	30%	15%
Low Value Items	50%	5%

The same degree of Control is not justified for all items. The class A items require the greatest attention and the Class C items the least attention. Class C items need no special calculation, since they represent a low inventory investment. The major concern of an ABC classification is to direct attention to those inventory items that represent the largest annual expenditures. If inventory levels can be reduced for class A items, a significant reduction in inventory investment will result. The ABC analysis is useful for any type of system (perpetual, periodic, optional replenishment and so forth).

Class A items deserve close control. Because each items represents a significant amount of inventory value, the perpetual system is often used since it provides the closest individual control. If. Class B items are of less value than class A items, and the optional replenishment system is frequently employed, It is usually more economical to combine orders tor several items. Class C items may have a policy with 6 month to 1 year purchase quantities. A two bin inventory system can be effectively used on these items, since at requires very little paper work. Class A and Class B items usually have a lower service level determined by an economic cost analysis. The advantage of an ABC analysis lies in a relaxing rather than in a tightening of inventory control (separate the 'Vital few' from the 'trivial many'). The relaxation of control comes from less emphasis put on C items, which represent the bulk of the inventory items. The ABC analysis gives a measure of inventory importance to each item.

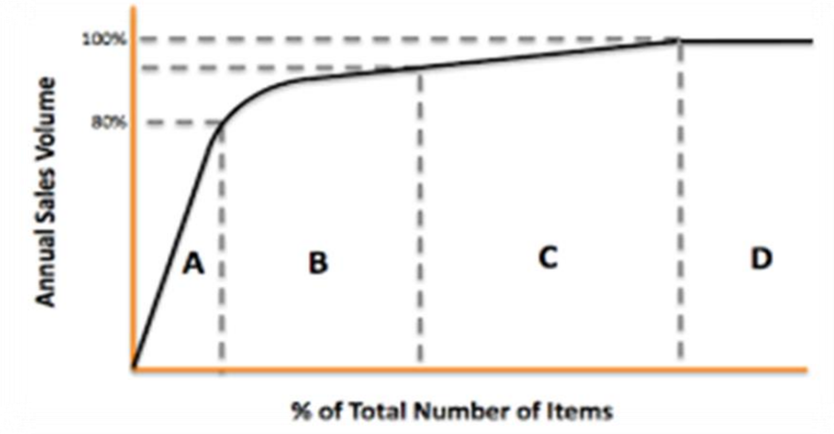


Fig.7: ABC Classification (Pareto Principle)

<u>The control policies for A. B. C items are based on the following principles:</u>

- ✓ To keep capital tied up in inventories as low as possible.
- ✓ To ensure that all the materials would be available when required.
- ✓ ABC analysis can be effectively used in material management. The various stages where it can be used are as follows:
  - o Information of items which require higher degree of control.
  - o To evolve useful recording strategy
  - o Stock record
  - o Priority treatment to different items
  - o Determination of safety stock items
  - o Stores layout
  - o Value analysis

The FSN Classification

This classification is done on the basis of the movement of materials from the store. Materials are classified as fast moving. Slow moving, and Non-moving (FSN), it is done on the basis of consumption pattern of items. This is required to control obsolescence. Cut off points of these three classes will depend on the peculiarities of the organization.

The SDE Classification

The materials are classified on the basis of problems faced in procurement. Here, the materials are sorted as Scarce to obtain, difficult to obtain and Easy to obtain. This classification is necessary for lead-time analysis and purchasing strategies. The tendency will be to procure more of scarce items. The next importance will be on items which are difficult to obtain. The least important are the items which are easy to obtain.

THE XYZ CLASSIFICATION

The classification is based on the value of items in storage. The items having high inventory values are X items. Low inventory value items are Z items. Y items lie between these two categories. This classification helps m finding out the items which are being excessively stocked. The main use is to review the inventories and their use at scheduled intervals.

The HML Classification

The classification is based on unit price of the material. In terms of high unit prices are High valued items, items of medium unit prices are Medium valued items, and items of low unit price are Low valued items. The cut off points will depend on the individual units. This is used mainly to control purchases.

1.6.2 Demand Management

A big data approach to decision making for logistics planning and scheduling using data collected via the use of RFID or sensors on manufacturing shop floors can be employed to assess customer demand and needs in real time (Zhong et al. 2015). Moreover, demand management (along with new manufacturing techniques and big data) can enable the supply chain to run concurrently, ensuring lower inventory costs and fast customer response times. Demand management also deals with the obsolescence of goods and wastage, especially for perishable goods, for which both agile and lean methodologies may be used (Christopher and Ryals 2014). To understand customers, global positioning system (GPS)-based surveys impose a lower respondent burden, offer greater accuracy and precision and incur fewer monetary costs, ultimately helping to understand demand and related issues better (Shankar 2015).

1.6.3 Bullwhip effect

The bullwhip effect refers to the gap that is caused by an increase in demand and a decrease in the volatility of inventories. The compelling study of the trade-off between responsiveness to demand and volatility can reduce the bullwhip effect. The bullwhip effect affects the SCM process through variations in a customer’s demand pattern and this result is amplified progressing through the production, supply and distribution processes. The impact of that distortion on the fill rate, inventory costs and transportation costs can be illustrated using a dynamics emulation model for the management of demand in multilevel supply chains (Bolarín, Frutos, and McDonnell 2009). Using agile methodology in high demand scenarios can prove beneficial for a firm's SCM (Lin and Lin 2006).

1.6.4 Total quality management and logistics

The three fundamental stages of supply chain procurement, production and distribution are undergoing transformation to keep up with market globalization and competitive pressure, as well as to ensure a quick response to customer needs. Competitive pressure forces firms to reduce costs and improve customer service with the help of IT and logistics options. Proper coordination between the various stages of a supply chain ensures a near-perfect supply chain model (Thomas and Griffin 1996). Accurately forecasting customer demand is a crucial part of providing a high-quality service, ultimately also leading to a positive impact on vendor rating. Goldman, Nagel, and Preiss (1995) noted that the customer's needs should be included in the development of product, process and service, further emphasizing the role of total quality management practices in relation to customer satisfaction.

1.6.5 Build-to-order supply chains

A build-to-order supply chain strategy helps improve the competitiveness of an organization by meeting the demands of individual customers through leveraging the benefits of outsourcing and IT. IT is one of the essential factors in enhancing operations and facilitating the implementation of build-to-order supply chains (Gunasekaran and Ngai 2005). Howells (2014) has mentioned that the future of supply chains is not that there will be only chains or no chains at all, but that they will transform into demand networks.

Chapter 2: BIG DATA ANALYTICS

2. Big data

Big data refers to the large and complex data assets that require cost-effective management and analysis for extraction of insights from them (Gupta & George, 2016). Four specific features, also known as the 4Vs, characterize big data (Kietzmann, Paschen, & Treen, 2018; Sivarajah, Kamal, Irani, & Weerakkody, 2017):

1. **Volume**: Volume refers to the large scale of big data, which requires innovative tools for their collection, storage, and analysis.

2. **Velocity**: Velocity refers to the rate at which the data are generated or updated, pointing to the real-time nature of big data.

3. **Variety**: Variety refers to the variation in types of data. Big data can come in diverse and dissimilar forms from multiple sources, such as texts, spreadsheets, audios, videos, and sensors. Big data are usually unstructured (e.g., text, audio) and are not organized in a structured manner in a relational data- base (e.g., tables, spreadsheets).

4. **Veracity**: Veracity refers to the complex structures of big data assets that make them ambiguous, imprecise and inconsistent. For example, the data related to consumer opinions posted on social media can be biased, inaccurate, and ambiguous.

5. **Value:** Value refers to the worth of the data being extracted. Having endless amounts of data is one thing, but unless it can be turned into value it is useless. While there is a clear link between data and insights, this does not always mean there is value in Big Data. The most important part of embarking on a big data initiative is to understand the costs and benefits of collecting and analysing the data to ensure that ultimately the data that is reaped can be monetized.

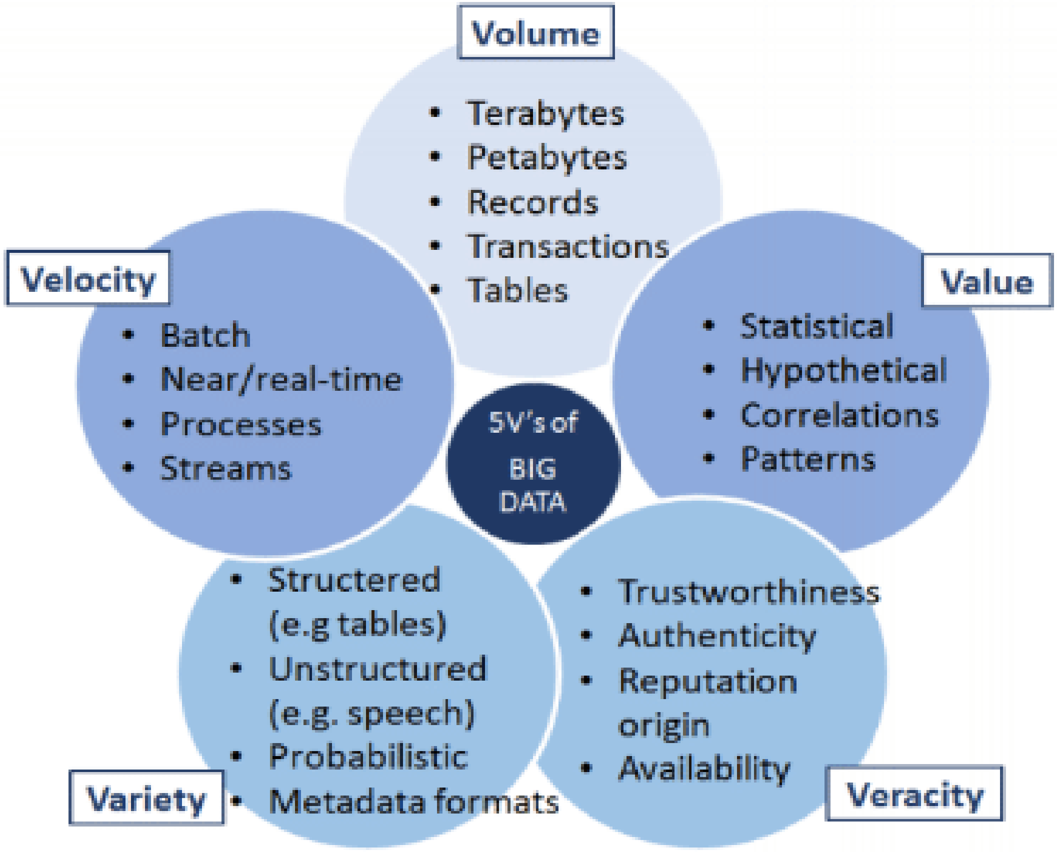


Fig 8: 5V characteristics of Big Data

2.1 Big data analytics cycle

The goal of big data analytics is to enhance organizational decision making and decision execution processes. Informed decision making is one of the building blocks of organizational success, and the importance of comprehensive analysis of information before making operational and strategic decisions has been highlighted in the works of many organizational researchers and practitioners (e.g., Dean & Sharfman, 1996; Fredrickson, 1984). In making important decisions, managers collect data, generate several alternative strategies, and carefully evaluate these strategies and their outcomes before making final decisions. Once implemented, the realized outcomes of the decision will be assessed to generate additional information that is cycled back into the subsequent decision making phases. The process of big data analytics resembles the aforementioned comprehensive decision making scheme used to enhance decision quality and out-comes. The big data analytics cycle is comprised of four important phases depicted in below figure:

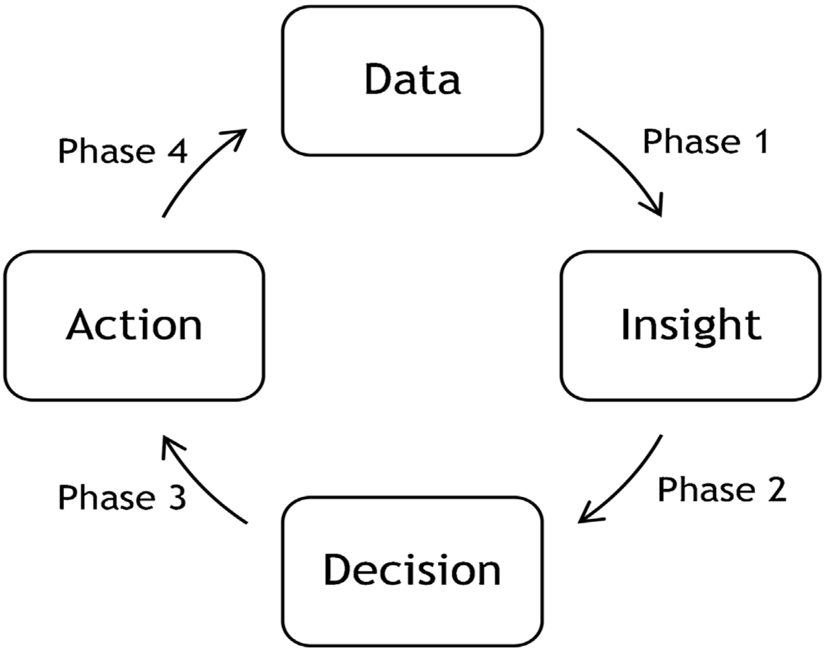


Figure 9: Big data analytics cycle

Phase 1: Large, diverse, and usually unstructured data are collected from internal and external sources and processed (i.e., cleaned and analysed) using advanced analytics tools and algorithms in order to generate insights. These insights are then interpreted by decision makers and used in the process of decision making.

Phase 2: Insights generated in Phase 1 are trans-formed into decisions. This is done by managers who contextualize the insights generated from their data analysis and attach meaning to them (Zeng & Glaister, 2017).

Phase 3: The decisions are transformed into specific operational actions. In other words, decisions are executed.

Phase 4: Transformation of decisions into actions generates additional outcomes (i.e., data points) which are cycled back into the process for future decision making efforts.

This way, a self-perpetuating cycle of big data analytics can significantly benefit organizational decision making processes and outcomes. Large volumes of both the internal operations data (e.g., inventory updates, employee performances, financial transactions, and consumer behaviours, sales) and the data collected from external sources (e.g., customer ratings, e-commerce communications, social media) are gathered and transformed into actionable insights.

2.2 BIG DATA ANALYTICS TOOLS

To present a brief and yet comprehensive account of big data analytics techniques, we classify these techniques into two broad categories: descriptive and prescriptive analytics tools (Sivarajah et al., 2017). For each category, we introduce several important artificial intelligence-based algorithms to clarify their applications. Table 3 summarizes the applications for each of the algorithms and provides specific real-world examples for each of them.

2.3 5 Types of Big Data Analytics

2.3.5.1 Prescriptive Analytics

The most valuable and most underused big data analytics technique, prescriptive analytics gives you a laser-like focus to answer a specific question. It helps to determine the best solution among a variety of choices, given the known parameters and suggests options for how to take advantage of a future opportunity or mitigate a future risk. Examples of prescriptive analytics for customer retention include next best action and next best offer analysis.

- Forward looking
- Focused on optimal decisions for future situations
- Simple rules to complex models that are applied on an automated or programmatic basis
- Discrete prediction of individual data set members based on similarities and differences
- Optimization and decision rules for future events

2.3.5.2 Diagnostic Analytics

Data scientists turn to this technique when trying to determine why something happened. It is useful when researching leading churn indicators and usage trends amongst your most loyal customers. Examples of diagnostic analytics include churn reason analysis and customer health score analysis. Key points:

- Backward looking
- Focused on causal relationships and sequences
- Relative ranking of dimensions/variable based on inferred explanatory power)
- Target/dependent variable with independent variables/dimensions
- Includes both frequents and Bayesian causal inferential analyses

2.3.5.3 Descriptive Analytics

This technique is the most time-intensive and often produces the least value; however, it is useful for uncovering patterns within a certain segment of customers. Descriptive analytics provide insight into what has happened historically and will provide you with trends to dig into in more detail. Examples of descriptive analytics include summary statistics, clustering and association rules used in market basket analysis. Key points:

- Backward looking
- Focused on descriptions and comparisons
- Pattern detection and descriptions
- MECE (mutually exclusive and collectively exhaustive) categorization
- Category development based on similarities and differences (segmentation)

2.3.5.4 Predictive Analytics

The most commonly used technique; predictive analytics use models to forecast what might happen in specific scenarios. Examples of predictive analytics include next best offers, churn risk and renewal risk analysis.

- Forward looking
- Focused on non-discrete predictions of future states, relationship, and patterns
- Description of prediction result set probability distributions and likelihoods
- Model application
- Non-discrete forecasting (forecasts communicated in probability distributions)

2.3.5.5 Outcome Analytics

Also referred to as consumption analytics, this technique provides insight into customer behaviour that drives specific outcomes. This analysis is meant to help you know your customers better and learn how they are interacting with your products and services.

- Backward looking, Real-time and Forward looking
- Focused on consumption patterns and associated business outcomes
- Description of usage thresholds
- Model application

Table 3: An overview of important big data analytics algorithms and applications

Categories	Algorithms Main Applications	Real-world examples
	Clustering Segmentation (e.g., customers, employees, products,	JPMorgan Chase uses clustering to segment millions of customers into different groups with different spending habits based on a combination of characteristics such as

Categories	Algorithms Main Applications	Real-world examples
Descriptive Tools	services)	transaction types and account balances. This can lead to devising effective strategies for delivering the best set of products to its customers.
	Association Rule Discovery Bundling (e.g., products, services, customers) for different purposes	Large supermarkets use this technique. In this regard, association rule discovery can reveal that, if customers buy onions, there is an 80% probability that they will also buy tomatoes. Such insight, derived from the analysis of billions of data points, enables Walmart store managers to quickly determine how to stock store shelves.
	Sequential Patterns Discovery Recommendation systems	Pharmacies use this technique to identify the temporal relationships between prescribed medications. The insights derived from such approach can trigger patient-specific recommendations. As another example, Netflix and Amazon Prime can analyze a user's previous choices of movies to recommend a set of movies and sort them based on the predicted preference of the user.
Predictive Tools	Classification Prediction of the class or group to which a new entity (e.g., customer or product) belongs	Large banks can rely on a classification model to predict if an existing customer is likely to open a new savings account. Such a predictive model, built through the analysis of behaviours (e.g., habits and transactions) and characteristics (e.g., age and gender) of customers, may enable the bank to identify easy-to-convert customers and engage in targeted advertising activities to attract them.
	Regression Prediction of a variable of interest (e.g., sales price,	Regression is heavily used in many industries. In the retail industry, for example, regression models predict future sales based on market conditions and other relevant historical data of

Categories	Algorithms Main Applications	Real-world examples
	consumer behaviour)	customers and competitors. Similarly, sophisticated regression models can be used to estimate real estate prices based on analysis of various data points such as time of the year and interest rates.
	Anomaly Detection Identifying outliers in data	Deutsche Bank utilizes an anomaly detection tool to discover fraudulent activity at the point of transaction. As another example, Intel uses image-based anomaly detection tools for the purpose of manufacturing quality control in its microchip production and assembly lines.

2.4 Data preparation (A mixture of art and science)

Basic data preparation operations access, transform, and condition data to create a dataset in the proper format suitable for analytical modelling. The major problem with data extracted from databases is that the underlying structure of the data set is not compatible with most statistical and data mining algorithms. Most data in databases are stored at the account level, often in a series of time-stamped activities (e.g., sales). One of the greatest challenges is rearranging these data to express responses on the basis of the entity to be modelled. For example, customer sales records must be gathered together in the same row of a data set for each customer. Additional preparation must be done to condition the data set to fit the input requirements of the modelling algorithm. There are a number of basic issues that must be addressed in this process.

Basic issues that must be resolved in data preparation:

- How do I clean up the data?—Data Cleansing
- How do I express data variables?—Data Transformation
- How do I handle missing values?—Data Imputation
- Are all cases treated the same?—Data Weighting and Balancing
- What do I do about outliers and other unwanted data?—Data Filtering
- How do I handle temporal (time-series) data?—Data Abstraction
- Can I reduce the amount of data to use?—Data Reduction
- Records?—Data Sampling
- Variables?—Dimensionality Reduction
- Values?—Data Discretization
- Can I create some new variables?—Data Derivation

2.5 Important Data Mining Techniques

One of the most important tasks in Data Mining is to select the correct data mining technique. Data mining technique has to be chosen based on the type of business and the type of problem your business faces. A generalized approach has to be used to improve the accuracy and cost-effectiveness of using data mining techniques. There are basically seven main Data Mining techniques which are discussed in this article. There are also a lot of other Data mining techniques but these seven are considered more frequently used by business people.

- Statistics
- Clustering
- Visualization
- Decision Tree
- Association Rules
- Neural Networks
- Classification

Statistical Techniques

Data mining techniques statistics is a branch of mathematics which relates to the collection and description of data. The statistical technique is not considered as a data mining technique by many analysts. But still, it helps to discover the patterns and build predictive models. For this reason, data analyst should possess some knowledge about the different statistical techniques. In today's world, people have to deal with a large amount of data and derive important patterns from it. Statistics can help you to a greater extent to get answers for questions about their data like

- What are the patterns in their database?
- What is the probability of an event to occur?
- Which patterns are more useful to the business?
- What is the high-level summary that can give you a detailed view of what is there in the database?

Statistics not only answer these questions they help in summarizing the data and count it. It also helps in providing information about the data with ease. Through statistical reports, people can make smart decisions. There are different forms of statistics but the most important and useful technique is the collection and counting of data. There are a lot of ways to collect data like Histogram, Mean, Median, Mode, Variance, Max, Min and Linear Regression.

Clustering Technique

Clustering is one of the oldest techniques used in Data Mining. Clustering analysis is the process of identifying data that are similar to each other. This will help to understand the differences and similarities between the data. This is sometimes called segmentation and helps the users to understand what is going on within the database. For example, an insurance company can group its customers based on their income, age, nature of policy and type of claims.

There are different types of clustering methods like Partitioning Methods, Hierarchical Agglomerative methods, Density-Based Methods, Grid-Based Methods, Model-Based Methods.

The most popular clustering algorithm is the Nearest Neighbour. The nearest neighbour technique is very similar to clustering. It is a prediction technique where in order to predict what an estimated value is in one record look for records with similar estimated values in a historical database and use the prediction value from the record which is near to the unclassified record. This technique simply states that the objects which are closer to each other will have similar prediction values. Through this method, you can easily predict the values of the nearest objects very easily. Nearest Neighbour is the easiest to use the technique because they work as per the thought of the people. They also work very well in terms of automation.

Visualization

Visualization is the most useful technique which is used to discover data patterns. This technique is used at the beginning of the Data Mining process. Many types of research are going on these days to produce an interesting projection of databases, which is called Projection Pursuit. There is a lot of data mining technique which will produce useful patterns for good data. But visualization is a technique which converts Poor data into good data letting different kinds of Data Mining methods to be used in discovering hidden patterns.

Induction Decision Tree Technique

A decision tree is a predictive model and the name itself implies that it looks like a tree. In this technique, each branch of the tree is viewed as a classification question and the leaves of the trees are considered as partitions of the dataset related to that particular classification. This technique can be used for exploration analysis, data pre-processing and prediction work.

The decision tree can be considered as a segmentation of the original dataset where segmentation is done for a particular reason. Each data that comes under a segment has some similarities in their information being predicted. Decision trees provide results that can be easily understood by the user.

CART which stands for Classification and Regression Trees is a data exploration and prediction algorithm which picks the questions in a more complex way. It tries them all and then selects one best question which is used to split the data into two or more segments. After deciding on the segments it again asks questions on each of the new segment individually.

Neural Network

Neural Network is another important technique used by people these days. This technique is most often used in the starting stages of the data mining technology. The artificial neural network was formed out of the community of Artificial intelligence.

Neural networks are very easy to use as they are automated to a particular extent and because of this the user is not expected to have much knowledge about the work or database. But to make the neural network work efficiently you need to know

- How the nodes are connected?
- How many processing units to be used?
- When should the training process be stopped?

There are two main parts of this technique – the node and the link

- The node **–** which freely matches to the neuron in the human brain
- The link – which freely matches to the connections between the neurons in the human brain

A neural network is a collection of interconnected neurons which could form a single layer or multiple layers. The formation of neurons and their interconnections are called the architecture of the network. There are a wide variety of neural network models and each model has its own advantages and disadvantages. Every neural network model has different architectures and these architectures use different learning procedures.

Association Rule Technique

This technique helps to find the association between two or more items. It helps to know the relations between the different variables in databases. It discovers the hidden patterns in the data sets which is used to identify the variables and the frequent occurrence of different variables that appear with the highest frequencies.

Association rule offers two major information-

- Support – How often is the rule applied?
- Confidence – How often the rule is correct?

This technique follows a two-step process

- Find all the frequently occurring data sets
- Create strong association rules from the frequent data sets

There are three types of association rule. They are

- Multilevel Association Rule
- Multidimensional Association Rule
- Quantitative Association Rule

This technique is most often used in the retail industry to find patterns in sales. This will help increase the conversion rate and thus increases profit.

Classification

Data mining techniques classification is the most commonly used data mining technique which contains a set of pre-classified samples to create a model which can classify the large set of data. This technique helps in deriving important information about data and metadata (data about data). This technique is closely related to the cluster analysis technique and it uses the decision tree or neural network system. There are two main processes involved in this technique

- **Learning** – In this process the data are analysed by the classification algorithm
- **Classification** – In this process, the data is used to measure the precision of the classification rules

There are different types of classification models. They are as follows

- Classification by decision tree induction
- Bayesian Classification
- Neural Networks

- Support Vector Machines (SVM)
- Classification Based on Associations.

2.6 Applications of Big Data

The application of big data analytics in SCM has been referred to as SCM data science (Waller and Fawcett, 2013), which includes the application of advanced quantitative and qualitative analysis to a vast volume of structured and unstructured data. Such analysis include predictive analytics (Schoenherr and Speier-Pero 2015), business analytics, big data analytics and supply chain analytics (Wang etal.2016).Predictive analytics, in particular, is a major factor in SCM in forecasting business trends and anticipated demand, minimizing stock-outs, even during periods of unanticipated demand, as in recent years. It can be used to understand the hidden potential of SCM concerning the skills required (Schoenherr and Speier Pero2015).Bigdataanalyticscanbeusedtostrengthenmarketcompetitiveness and improve data quality management and the data usage experience. A positive relationship has also been shown between maintaining the quality of big data and the perceptions of firms towards adopting big data analytics via internal or external sourcing of data (Kwon, Lee, and Shin 2014). In spite of the many benefits of the application of big data in supply chains, there are certain barriers to implementing predictive analytics, such as the lack of skilled professionals, lack of International journal of logistics research and applications awareness and a dearth of tools for training the next generation of data scientists in the supply chain industry (Schoenherr and Speier-Pero 2015).

Predictive analytics and forecasting

Predictive analytics of sales data can be used to predict and forecast future demand for goods (Chase 2015). Predictive analytics is also used to examine customer purchase behaviour and provide purchase suggestions (Greco and Aiss 2015). The use of sensory networks to predict the remaining life of perishable goods is also possible with the help of predictive analytics (Li and Wang 2015).

Tracking of goods

Tracking tools that use sensory and tracing data provide support for supply chain decision making concerning logistics and SCM. Sensory data-driven supply chains exhibit better pricing models and better performance (Li and Wang 2015). According to Gartner (2014), to enhance in-transit visibility, IoT will ‘significantly alter how the supply chain operates’ and specifically the impact will relate to ‘how supply chain leaders access information’.

Vendor-managed inventory and centralized planning and forecasting

Lack of collaborative practices among vendors causes challenges in inventory management, especially for perishable goods. Centralized forecasting and planning based on retailers' sales data can be used to forecast the entire supply chain and ensure better responsiveness to demand through the improved availability of products (Alftan et al. 2015).

Use of analytics to improve accuracy

Problemforecastingcanbeusedtoaddresspotentialproblemsproactivelybeforetheyoccurthroug h predictive monitoring techniques, such as machine learning and constraint satisfaction using quality of service agreements. Predictive monitoring can help reduce lead times by 70%. Precision and accuracy can be improved up to 14% using constraint satisfaction through quality of service agreements. Moreover, there allratecanbeimprovedby23%usingmachinelearningandconstraintsatisfaction (Metzger et al. 2015).

2.7 DATA SCIENCE

Data science, predictive analytics and big data, known collectively as DPB, play a vital role in decision making. DPB also ensures competitiveness by assessing the past and future integration of the business processes, cost levels and service levels of companies (Waller and Fawcett 2013). Advanced analytics is likely to become a decisive competitive asset in many industries and is a core element in a company's efforts to improve performance using data science principles (Waller and Fawcett 2013). The use of RFID cuboids to establish data warehouses, mapping with other cuboids and using spatiotemporal sequential logistics trajectories to perform logistics operations are examples of the application of data sciences in the supply chain industry (Zhong et al. 2015).

Collaborative partnership to reduce risks

Corporate logistics is used to share information with the aim of enabling the integration of disparate information amongst supply chain partners. The sharing of strategic information and customizable information technology ensure performance gains and symmetry of participation (Klein, Rai, and Straub 2007). Investments in IT make the greatest contribution to competitiveness in enabling dynamic supply chain capabilities (Fawcett et al. 2011).

Lack of trust and competitive pressure

The importance of information in a downstream supply chain is considerable as inputs are proffered to those immediately upstream. Steps to minimize data distortion by competitors are of the utmost importance to prevent sales or order variance. Distrust among suppliers or vendors can have a negative impact on the four sources of the bullwhip effect, i.e. demand signal processing, the rationing game, order batching and price variations. Steps should be taken to ensure the accurate flow of information concerning delivery plans, production scheduling and inventory control (Lee, Padmanabhan, and Whang 2004).

2.8 R Programming Language

R is a programming language developed by Ross Ihaka and Robert Gentleman in 1993. R possesses an extensive catalogue of statistical and graphical methods. It includes machine learning algorithm, linear regression, time series and statistical inference to name a few. Most of the R libraries are written in R, but for heavy computational task, C, C++ and FORTRAN codes are preferred.

R is not only entrusted by academic, but many large companies also use R programming language, including Uber, Google, Airbnb, Facebook and so on.

Data analysis with R is done in a series of steps; programming, transforming, discovering, modelling and communicate the results

- **Program**: R is a clear and accessible programming tool
- **Transform**: R is made up of a collection of libraries designed specifically for data science
- **Discover**: Investigate the data, refine your hypothesis and analyze them
- **Model**: R provides a wide array of tools to capture the right model for your data
- **Communicate**: Integrate codes, graphs, and outputs to a report with R Markdown or build Shiny apps to share with the world.

2.9 Why use R?

Data science is shaping the way companies run their businesses. Without a doubt, staying away from Artificial Intelligence and Machine will lead the company to fail. The big question is which tool/language should you use?

They are plenty of tools available in the market to perform data analysis. Learning a new language requires some time investment. The picture below depicts the learning curve compared to the business capability a language offers. The negative relationship implies that there is no free lunch. If you want to give the best insight from the data, then you need to spend some time learning the appropriate tool, which is R.

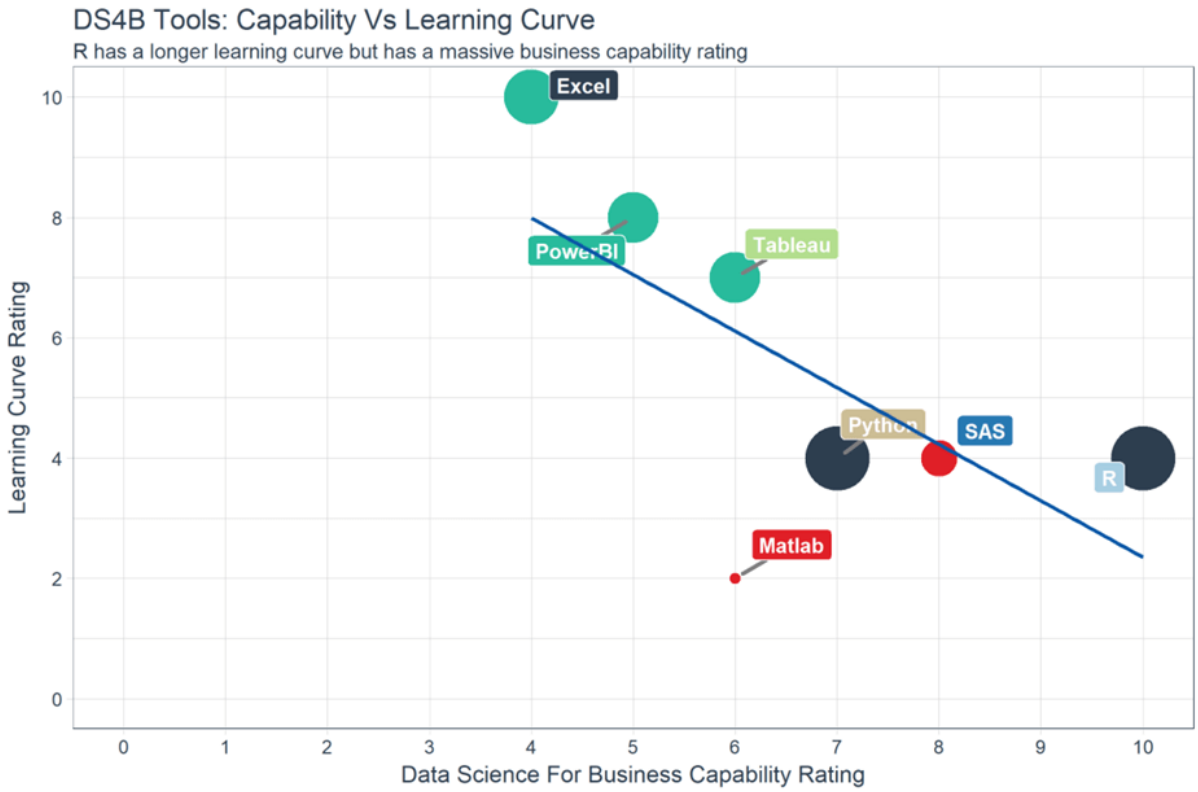


Fig.10: The learning curve compared to the business capability

On the top left of the graph, you can see Excel and PowerBI. These two tools are simple to learn but don't offer outstanding business capability, especially in term of modelling. In the middle, you can see Python and SAS. SAS is a dedicated tool to run a statistical analysis for business, but it is not free. SAS is a click and run software. Python, however, is a language with a monotonous learning curve. Python is a fantastic tool to deploy Machine Learning and AI but lacks communication features. With an identical learning curve, R is a good trade-off between implementation and data analysis.

When it comes to data visualization (DataViz), you'd probably heard about Tableau. Tableau is, without a doubt, a great tool to discover patterns through graphs and charts. Besides, learning Tableau is not time-consuming. One big problem with data visualization is you might end up never finding a pattern or just create plenty of useless charts. Tableau is a good tool for quick visualization of the data or Business Intelligence. When it comes to statistics and decision-making tool, R is more appropriate.

Stack Overflow is a big community for programming languages. If you have a coding issue or need to understand a model, Stack Overflow is here to help. Over the year, the percentage of question-views has increased sharply for R compared to the other languages. This trend is of course highly correlated with the booming age of data science but, it reflects the demand of R language for data science.

R is a great tool to explore and investigate the data. Elaborate analysis like clustering, correlation, and data reduction are done with R. This is the most crucial part, without a good feature engineering and model, the deployment of the machine learning will not give meaningful results.

Chapter 3

3.1 PRIOR ART

3.1.1 MULTI-CRITERIA INVENTORY CLASSIFICATION (MCIC) - BACKGROUND AND MOTIVATION

Because of its easy-to-implement nature, applicability to numerous situations, empirically observed benefits and remarkable effectiveness in many inventory systems, this approach is still popularly used in practice. However, the method has a serious drawback that may inhibit the effectiveness of the procedure in some situations. The criterion used in the conventional ABC classification is the annual dollar usage, so using one criterion may create problems of significant financial loss.

Hence, only one criterion is not always very efficient measure for decision making. Therefore, multiple criteria decision making (MCDM) methods are used (FIores, Whybark, 1986, 1987). Apart from annual dollar value, other criteria like lead time of supply, part criticality, availability, Stock-out penalty cost, ordering cost, scarcity, durability substitutability, reparability, risk of obsolescence, etc., have been taken into consideration (Flores, Whybark, 1986, 1987; Zhou and Fan, 2007).More studies have been carried-out on the field of MCIC in the past 20-25 years. So many different methods for classifying inventory and taking into consideration multiple criteria have been used and developed. The considerable body of research (as discussed in literature review section) shows many feasible ways of implementing multiple criteria ABC analysis in practice — some extremely simple while others quite sophisticated. The earlier researches based on Analytic Hierarchy Process (AHP) in this direction suffer by subjectivities associated with assignment of weights from expert judgment (Flores, Partovi, and Kabir). The unavoidable biasness and familiarity of different experts with each of the criterion may lead to inconsistencies and unsatisfactory results. While optimization techniques and soft computing based approaches (rough set theory, artificial intelligence, clustering, etc.) by large were motivated to get away from subjective weights, it is realized that subjectivity in this context is a good thing to reflect the management priorities. In some of the approaches even though management decisions were additionally captured in the form of exogenously ranking the weights by the Decision Maker (DM), the interpretation of What did not seem to be simple and comprehensive enough in the perspective of further training requirements with them.

The second point of concern is that the several previous methods in the literature are fully compensatory in multiple criteria aggregation (Liu, 2016). This means that an item scoring badly on one or more key criteria may still be placed in the best class because these bad performances could be compensated by other criteria. Thus, it is necessary to consider the non-compensation in the multiple criteria ABC analysis. To the best of our knowledge the ABC classification problem with non-compensation among criteria has not been studied sufficiently. Some exceptions include the studies developed by Zhou and Fan, Hadi- Vencheh, and Lolli et al. The third known issue is the limitation associated with imposing fixed cardinality on each classification. The limit of cardinality should rest solely on manager's decision taking into consideration the company's vision, service level requirements for customers etc.

The need of a combinatorial but transparent way of dealing with weights of criteria is addressed in this paper. This is accomplished by integrating subjective weights considering fuzzy environment, two types of objective weights based on the concept of information entropy and standard deviation and inter-attribute correlation effect based weights.

In the proposed approach, taking into account the requirement of the ABC analysis, the cardinality limitation of items in each class is specified in advance (A-20%, B- 30%, C- 50%). Hence, the previous approaches which do not have the limitation in place are not compared with the proposed method.

The Multi Objective Optimization on the basis of Ratio Analysis (MOORA) and its updated form (MULTIMOORA) methods are effective and simple MCDM techniques. In this dissertation, the inventory classification is based on both MULTIMOORA and a comprehensive form of the same method which is referred to as EXTENDED MULTIMOORA method. The proposed methodology in solving the MCIC problem was developed through considering target-based normalization technique and integrated significant coefficients.

As one can note, the Reference Point prevents the MULTIMOORA from becoming a fully compensatory technique. Whereas the Ratio System and the Full Multiplicative Fonn are fully compensatory methods, the Reference Point is not of that kind. MULTIMOORA combines all these three subordinate methods and arrives at final ranking of the alternatives.

3.1.2 MCIC METHODS

Since Flores and Whybark (1987) first proposed looking at more than one criterion, this has been an area of active research. There has been broad agreement that ABC analysis should consider more than one criterion. The methodology involves three main steps once the relevant criteria have been identified. The first is to determine what weights to assign to the different criteria and the second is to score each item on each criterion. If the criteria are measured on a variety of scales, this second step might involve rescaling the scores onto a 0-1 or 0-100 scale. The final step is to combine weights and scores to produce the weighted score. Over the years, three broad approaches have emerged to perform the weighting. It has been assumed that the different criteria permit unambiguous scoring of the items and that this is not an issue.

3.1.3 SUBJECTIVE WEIGHTING AND RATING

This approach scores each type of inventory item on each criterion and then combines the different scores using a subjective weighting scheme. Many researchers have used the framework provided by the AHP to accomplish this (Flores, Olsen, & Dorai, 1992; Partovi & Burton, 1993; Partovi & Hopton, 1994; Gajpal, Ganesh, & Rajendran, 1994; Kabir, Hasin, & Khondokar, 2011; Braglia, Grassi, & Montanari, 2004). AHP relies on pairwise comparisons of criteria with respect to an overall objective to derive the weights for the criteria. Alternatives too can be compared pairwise with respect to each criterion. In this case, the alternatives are the various inventory items. Pairwise comparison of thousands of items with respect to each criterion is clearly a mammoth task. Instead, the alternatives are rated along each criterion and the weights are applied to these ratings. This is AHP in its ratings mode. The result is a weighted rating that can be used to rank the items prior to classifying them into different categories. The pairwise comparisons needed to determine the weights are performed by managers who are knowledgeable about the inventory items and the trade-offs among the different criteria to deal with. This is a one-time task as long as the set of criteria or the management preferences among them do not alter.

AHP has found place in a variety of business decision-making phases and decision-makers also have found it intuitive and easy to use (Saaty, 1995; Zahedi, 1986; Vargas, 1990). Its theoretical underpinnings are strong and it has been incorporated into software (Expert Choice) that makes the implementation of the decision making process easy.

There are alternative ways of implementing rating and weighting schemes. Researchers might not have indicated them in the light of ABC analysis. For example, Multi-Attribute Utility

Theory provides theory and methodology for assessing weights, rating alternatives, and combining weights and ratings to arrive at a final rating (or utility) for an alternative. The most robust and easy to use model is an additive model that is very similar to the AHP in its ratings mode. Software, as for example, SMART (Edwards & Barron, 1994) [**16**] also exists by which this process can be easily implemented.

Whichever method is opted for use, once the weights are obtained, the weighting and rating can be simply performed on a spreadsheet.

3.1.4 LINEAR OPTIMIZATION

Other researchers (Ramanathan, 2006; Ng, 2007; Zhou & Fan, 2007; Hadi-Vencheh, 2010) utilized a linear optimization approach to determining the weights. Their view is that the subjective inputs needed in the weighting and rating approach are cumbersome to obtain and undesirable because of possible inconsistencies. Instead, they would rather let the data itself decide weights that minimize some reasonable criteria or objective functions.

Ramanathan (2006) solved a linear programming problem for each item in inventory to determine weights that maximize the weighted score for that item subject to constraints that the weighted sum for every item using this same set of weights is less than or equal to one. Thus, one immediate criticism of this model is that with more than a handful of items, the process will not only become more cumbersome but would also consume more time.

Ng (2007) addressed this issue by proposing a DEA-type model similar to Ramanathan's. By this technique, the original optimization model is then transformed into another set of problems, the structure of which makes it easy to recognize the optimal solution without the use of a linear optimizer. Apart from endogenous derivation of final criteria weights by the optimal solver, the decision makers are also allowed to exogenously specify the exact values of weights in the form of a ranking of the weights associated with the criteria for each item, but this ranking is not critical to the mechanics of the method which can be implemented on a spreadsheet. At the end of the process, each item in inventory is given a rating which could then be used to perform the ABC analysis. Hadi-Vencheh (2010) proposed a nonlinear extension to the Ng model.

A second criticism of Ramanathan's model is that the method can provide high scores to items that score highly on an unimportant criterion. Zhou & Fan (2007) proposed a refinement which avoids this problem.

3.1.5 CLUSTERING, GENETIC ALGORITHMS, AND NEURAL NETWORKS

A third approach to categorization for the purpose of ABC analysis relies on the methods of artificial intelligence and data-mining. All these methods start with a training set – a set of inventory items that have already been classified on the basis of multiple criteria as A, B, or C, by managers who are familiar with them - to learn the appropriate transformations necessary to combine criteria values and determine cut-offs.

Guvenir and Erel (1998) proposed an approach called GAMIC which starts with the framework of AHP to deal with multi-criteria ABC analysis. GAMIC uses genetic algorithms to learn from the training set the weights to be assigned to each criterion and, further, to determine the cut-offs between the three categories. Unknown weights and cutoffs are encoded as chromosome vectors that result in a particular classification. Given this encoding scheme, the method applies standard genetic operators (reproduction, crossover, and mutation) to create new generations of solutions. Each chromosome (solution) is tested using a fitness function and the best solutions become members of the next generation. This process continues iteratively until the algorithm converges on the training set; i.e., provides weights and cut-offs that reproduce (for the training set) the decision-maker's categorizations. These weights and cut-offs can then be used for other inventory categorization tasks. In their comparisons, their algorithm performed better than AHP – in the sense of having fewer misclassifications when compared with the decision-maker's classifications of the items. One limitation of this approach is that criteria can only be quantitative.

Partovi and Anandarajan (2001) followed a similar process but using Artificial Neural Networks (ANN) to solve an inventory classification problem with four criteria - unit price, ordering cost, demand range, and lead time. The inputs to the network are values of these criteria for different inventory items. The output of the network is a categorization of a set of criteria values as A, or B, or C. Thus, their network consists of four input neurons (one for each input criterion), 16 hidden neurons, and 3 output neurons (one for each inventory category). Two kinds of learning algorithms were used - back propagation and genetic algorithms. Once the network was trained, it was used on stored data (validation) as well as on test sample. Results (% misclassification compared with decision-maker categorization) were encouraging and proved ANN to be viable of performing multi-criteria ABC analysis. Yu, 2011 also used ANN for solving MCIC problem.

Gulsen and Ozkan (2013) treated ABC analysis as a clustering problem in which the inventory items that have to be categorized are partitioned into 3 "fuzzy" clusters by minimizing some appropriate clustering function. Fuzzy clustering is the appropriate technique to use given that it is possible for some inventory items to belong to more than one cluster. The centre of a cluster is described by an n-dimensional vector, where n is the number of criteria to be used for the ABC analysis. Each inventory item is similarly an n-dimensional vector. Membership of the clusters is indicated by a membership value that is between 0 and 1. The objective to be minimized is the distance between the current centres of each cluster and each inventory item weighted by the membership value modified by a "fuzzifier". The algorithm starts with initial values for the cluster centres, followed by calculating a membership value for each inventory item. This allows recalculation of the cluster centres. If the new cluster centres are within some ε of the current cluster centres, the algorithm stops; otherwise, the next iteration begins with the new cluster centres. Once the stopping rule has been met, the output of the algorithm is the membership value for each item for each cluster. An item is assigned to a cluster based upon the highest of its membership values. Thus, at the end of the process, three (for three categories) clusters will have been identified. The next step is to label the clusters appropriately. Labelling is done on the basis of the average criterion value within a cluster. This is calculated by adding all the criterion values for all items within a cluster and dividing by the number of items in the cluster. The cluster with the highest average criterion value is labelled A, the next highest as B, and the last one as C. In actual application of the method, it is suggested that item ratings on each criterion be rescaled to a 0-1 scale using a simple linear transform.

In concept, each of the above three approaches will produce an ABC categorization with high reliability; in other words, there is a high degree of overlap with the categorizations of human decision-makers.

Tsai and Yeh proposed a particle swarm optimization approach for the multi-criteria ABC problem.

Lolli et al. (2012) [**32**] used a k-Means clustering for solving the MCIC problem, which is very similar to above described method, then it solves the compensation issue by utilizing a veto rule on top of the proposed algorithm.

3.1.6 OTHER APPROACHES

Other approaches have been proposed to the ABC categorization problem. Rough set theory (Pawlak, 1991) has been used by Gomes and Ferreira (1995) and Chen, Li, Levy, Hipel, and Kilgour (2008) to perform the ABC categorization with the use of training sets. Bhattacharya, Sarkar, and Mukherjee (2007) presented a distance-based consensus method using the concepts of ideal and negative ideal solutions from the TOPSIS (Technique for Order

Preference by Similarity to Ideal Solution) approach to ranking. They demonstrate the practicality of their approach by applying it to the inventory items of a pharmaceutical company. Liu & Huang (2006) and Torabi, Hatefi, & Pay (2012) presented modified versions of a DEA model to take both quantitative and qualitative criteria into account in ABC analysis.

MOORA

In one way or the other, MCIC is typically a multi-criteria decision making problem. This involves several criteria or factors on which decision maker's knowledge is usually vague and imprecise.

The MULTIMOORA method (Brauers and Zavaskas 2010) is a recently introduced new MCDM method based on multi-objective optimization by ratio analysis. (MOORA) (Brauers and Zavaskas, 2006). Due to its characteristics and capabilities, the use of MULTIMOORA has been increasing in the literature. Brauerrs et al. (2013) employed the MULTIMOORA to analyze the construction sectors of European countries from a macroeconomic point of view by comparing construction market variations appeared during recession. Streimikiene and Balezenitis (2013) proposed a MCDM methodology by using the MULTIMOORA for assessing mitigation strategies for climate changes and applied it for determining the optimal mitigation policies in Lithuania. Brauers et al. (2012) used the MULTIMOORA to estimate economic worth of European Union (EU) member states towards 2020. Karande and Chakraborty (2012) used the ratio system, the reference point approach, and the full multiplicative form to solve some of the common material selection problems. Mandal and Sarkar (2012) used the fuzzy MULTIMORRA for comparing the intelligent systems of conflicting nature. Lie et al. (2014) used the MULTIMOORA method under fuzzy environment to evaluate the failure modes to address the problem of infant abduction from hospital. Mishra et al. (2015) applied fuzzy integrated MULTIMOORA method towards supplier/partner selection in agile supply chain. Hafezalkotob and Hafezalkotob (2015) successfully used the target-based MULTIMOORA with integrated significant coefficients is solving material selection in biomedical engineering. Dey et al. presented a supplier selection strategy selection methodology based on fuzzy MULTIMOORA.

3.1.7 TOPSIS METHOD

TOPSIS is a well-known MCDM method, which was developed by Hwang and Yoon in 1981. It helps decision-makers to rank a set of alternatives evaluated on a collection of conflicting and non-commensurable criteria. The basic idea of this method is to rank first the alternatives which have the shortest distance from the Positive Ideal Solution (PIS) and the farthest distance from the Negative Ideal Solution (NIS).

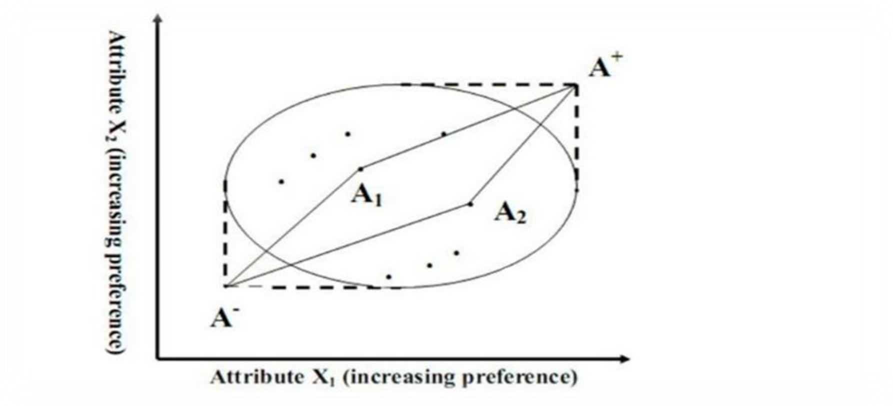


Fig 11: Basic concept of TOPSIS method (A+: Ideal point, A−: Negative—Ideal Point).

The PIS (resp. NIS) is a ”hypothetical alternative” that has the best (resp. worst) values for all considered criteria. Formally, for a problem with N alternatives Ai (i =1... N) evaluated on M criteria Cj (j=1,...,M), the main steps of TOPSIS method are illustrated as follows:

Step 1: Construct the performance matrix $X = (x_{ij})_{N,M}$in which each alternative

Ai (i=1... N) is evaluated on the criterion Cj (j=1,...,M).

Step 2: Determine the criteria weights w_j(j=1,..., M) such that:

$$\sum_{j=1}^{M} w_j = 1$$

Step 3: Compute the normalized decision matrix$R = (x_{ij}^n)$):

$$x_{ij}^n = x_{ij} / \sqrt{(\sum_{k=1}^{N} x_{kj}^2)} \text{ ; j = 1... M, i = 1, ..., N.}$$

Step 4: Compute the normalized weighted decision matrix$V = (v_{ij})_{N,M}$:

$$v_{ij} = w_j x_{ij}^n;\ j = 1, \ldots, M, i = 1, \ldots, N.$$

Step 5: Determine the PIS and NIS:

$$PIS = A_i^+ = \{V_1^+, V_2^+, \ldots\ldots, V_m^+\} = \{(\max\{v_{ij}\}|j \in B, min\{v_{ij}\}|j \in C)\}$$

$$NIS = A_i^- = \{V_1^-, V_2^-, \ldots\ldots, V_m^-\} = \{\left(min\{v_{ij}\}\middle| j \in B, max\{v_{ij}\}\middle| j \in C\right)\}.$$

Where B and C are respectively the benefit and the cost criteria sets.

Step 6: Compute the Euclidean distance S_i^+ $(resp. S_i^-)$between each alternative A_i and PIS (resp. NIS), let:

$$S_i^+ = \sqrt{\sum_{j=1}^{M} (v_{ij} - V_j^+)^2} ; i = 1, \dots., N$$

$$S_i^- = \sqrt{\sum_{j=1}^{M} (v_{ij} - V_j^-)^2}; i = 1,2, \dots, N$$

Step 7: Compute the relative separation measure (or a score) of each alternative Ai :

$$SM_i = \frac{S_i^-}{S_i^+ + S_i^-} ; i = 1,2, \dots., N$$

Finally, the alternatives are ranked in a descending order according to theirSM_i and a category is then assigned to each alternative.

3.2 Gap Analysis:

A research gap analysis should be used to analyze gaps in research processes and the gulf between the existing outcome and the desired outcome. This step process can be illustrated the example below:

Identify the existing process: climbing stairs to reach flat which is in twentieth floor

- Identify the existing outcome: Walking the stairs to reach his flat
- Somehow he/she can manage up to tenth floor by walking the stairs
- Identify the desired outcome: she/he wants to reach his flat
- Identify the process to achieve the desired outcome: she/he can use an alternative method such as using the lift
- Identify and document the gap: it is a difference of ten more floor
- Develop the means to fill the gap: she/he acquires and uses the lift
- Develop and prioritize requirements to bridge the gap.

Traditional paradigm for supply-chain management is to develop sophisticated tools to generate forecasts that accurately predict the value and the level of uncertainty of future demand. These forecasts are then used as an input to an optimization problem that evaluates trade-offs and respects constraints in order to come up with decisions about managing materials. This two-step process, which is embodied in all current material-management planning and control systems, can be replaced by a single-step process that looks for the best relationship among all of the data and the decisions. Based on learning from the past, a “best” relationship can be identified, which will generate decisions, as future uncertainty is resolved, that are better than the decisions derived from the traditional two-step approach of first forecast and then optimize.

This approach is not restricted by any a priori assumptions about the nature of the market and the behaviours that lead to customer demands or about the trade-0ffs and constraints that have to be considered in order to evaluate material-management decisions. Instead, the power of computer learning, supplemented by management input based on context-specific knowledge, is used to find the best relationship between all possible decisions and full range of the data. Use of this relationship can lead to better operational performance. It will lead to better outcomes because it utilizes all of the data available to current methods along with extensive additional data that currently is ignored and which may be relevant.

This scenario becomes even more compelling when the impact of the internet of things is factored in. As they are used by customers, smart, connected products can generate orders of magnitude more data about current operating conditions and real-time performance of products. This data, along with traditional historical sales-based data, can support better methods for maintenance and the replacement of products.

The approach described here extends the concept of prescriptive analytics, which is considered by many to be the ultimate use of Big Data. Prescriptive analytics, however, has eluded most users of Big Data to date. There are some notable exceptions in industries such as online apparel retailing, where companies can view real-time, customer-purchase decisions (e.g., to buy or not to buy) and also can change the price of each product frequently at a negligible cost. The online retailer, however, knows little about the probability that consumers will purchase at each prices it sets but can learn dynamically about expected demand from sales data.

While many challenges remain, it is clear that a new approach that exploits all of the data that is becoming available is inevitable, given the connectivity, capacity, and transparency of data sources along with the vast computing power and data storage capacity available at a low cost. Like all planning systems, the proof will be in the results, when intelligent systems based on this approach are applied in practice. Change is coming to the world of inventory management and those that embrace this change will be ahead of the game. Successful adoption of this change will require active involvement of multiple functions within the firm along with a high level of coordination with both upstream and downstream supply chain partners as well as engagement with customers.

Table 3: Gap Analysis of MCIC & Big data Analytics in inventory management Literature

SL No	Technique	Author	Year	Areas covered	Issues not addressed
1	AHP	Flores et al	1992	Used AHP to reduce multiple criteria to single and consistent measure considering multi objectives	Qualitative criteria not considered. May suffer from subjective biasness and imprecise specifications.
2		Gajpal et al.	1994	Various modes of criteria considered in AHP and absolute measurement of part criticality achieved	Did not consider other classification techniques. May suffer from subjective biasness and imprecise specifications
3	TOPSIS	Bhattacharya et al	2007	Used TOPSIS model to classify items as A, B, and C	Fuzzy classifier not considered. May suffer from subjective biasness and imprecise specifications.
4	ABC Pareto	Reid	1987	Used ABC analysis for items in a Respiratory Therapeutic Unit	Only single criteria like annual dollar

SL No	Technique	Author	Year	Areas covered	Issues not addressed
					consumption value considered. Does not comply with managerial judgment of giving high importance to the slow moving but critical moderate value items.
5	Data Analytics	Waller and Fawcett	2013	Using quantitative and qualitative methods to improve supply chain design and competitive power by analysing past data and by integrating business processes, functions, costs and service levels with the help of big data analytics in SCM.	Managers need to understand and embrace the role of data science, predictive analytics and big data (DPB) and the implications for supply chain decision making.
6		Kwon, Lee, and Shin	2014	Use of big data analytics to strengthen market competition and to open up new business opportunities, internal or external sourcing of data, data quality management and data usage experience solutions.	Hesitance of firms in adopting big data.
7		Zhong et al.	2015	Implementing big data approach in decision making, i.e. logistics	Use of more sophisticated systems with

SL No	Technique	Author	Year	Areas covered	Issues not addressed
				planning and scheduling, with data collected using RFID on manufacturing shop floors. Mix of supply chain quality management and technology to improve SCM.	improved technologies to improve quality. Lack of information management in supply chain.
8		Da Xu, He, and Li	2014	Building powerful industrial systems and applications exploiting the ubiquity of RFID, wireless and mobile technology and sensors using IoT.	Implementation of major IoT applications in the industry was not mentioned.
9		Ng et al.	2015	Impact on SCM of the development of IoT or Internet-connected objects (ICOs) to meet customer needs by incorporating personal ICO data into various customisable applications (a 'platform strategy') and by maximising consumer value	Providers need to put mechanisms in place to enable customised solutions to emerge and place their strategic focus on their platform and the design of standardised interfaces. In case of demand management, Slow transition from product to

SL No	Technique	Author	Year	Areas covered	Issues not addressed
					platform focus, which requires a shift in supply chain logic from linear to network, web or eco-system thinking.
10	Cluster analysis	Canetta at al	2005	Used self-organizing map (SOM) based clustering approach.	Computationally expensive. Introduction of new items leads to re-clustering. Managerial preference cannot be combined.
11		Ernst and Cohen	1990	Used the concept of operation related groups (ORG) by means of statistical clustering utilizing full range of significant attributes.	Computationally expensive. Introduction of new items leads to re-clustering. Managerial preference cannot be combined.
12	Optimization	Zhou and Fan	2007	Extended Ramanathan's by most favourable and least favourable.	AHP not used
13		Hadi Vencheh	2010	Non-linear extension to Ng's model of Optimization.	Qualitative criteria not considered
14		Torabi et al	2012	Modified version DEA model using linear optimization for both quantitative and	The method does not consider integration with modem AI

SL No	Technique	Author	Year	Areas covered	Issues not addressed
				qualitative criteria	techniques. Fuzziness not considered
15	Neural networks	Huiskonen et all	2005	Used ANN for evaluating importance of C items for customer specific factors	ABC analysis in general was missing as it assumed as A and B class items correctly classified.
16		Partovi and Anandarajan	2002	Classified inventory items based on artificial neural networks modelling integrated with GA.	ANN is not hybridized with any MCDM technique. Fuzziness not considered. Prerequisite of training data already classified.

3.3 AIMS:

The final aim of any work in operations management is smooth, sustainable, and profitable operations. It has to take the comparative economic advantages by virtue of sensible management decisions in a highly competitive environment with an additional inclination towards better social, political, and environmental wellbeing. The inventory classification, which in broader sense often coined as Stock Keeping Unit (SKU) classification, is no exception to that.

SKU classification is a matter of importance for any business of any size from very large to very small. The huge body of research directed towards achieving a desired classification speaks out in this favour. For running a production process smoothly, there must be proper planning, hassle-free on-demand procurement, efficient and effective manufacturing, prompt delivery, and monitoring and controlling of the entire process for getting the desired customer satisfaction.

3.4 OBJECTIVES:

Traditional Inventory Management to Data driven Inventory Management

i. Traditional inventory systems have always focused on improving forecasting, while the forecast is only a small part of the overall inventory management problem. Most of the issues occur at the retail store shelf, which aren't included in forecasts.

ii. A new paradigm is needed – one which brings all factors into play, and which enables a proactive approach to solving inventory problems before they occur.

Lack of Visibility to Explicit Vision

i. The complexity of the data driven inventory management has increased with globalisation; accurate inventory has become a critical factor for success. Having complete viability and insight into your inventory items from supplier to customer is essential to effectively managing a global supply chain network.

ii. You need right items at right quantity at right time at reasonable cost. Questions are-

   a. How will you get them?
   b. When will you get them?
   c. How much cost involved in this process?

Failure to Keep Track of Stock to Meticulous Observation

Infrequent inventory checks and using manual processes to track inventory is not enough to manage your supply chain. Failing to keep close track of stock movement in and out of your business causes accounting errors, resulting in added costs for your business.

Low Product Turnover to Methodical Conceptualization

Failure to track demand can lead to a lower than expected product turnover due to low demand. Low turnover results in excess inventory that ends up wasting space in your warehouse and tying up capital, which is costly for business.

3.5 SCOPE OF PRESENT INVESTIGATION

Inventory data management deals with large collection stock related data in the supply chain management environment. The frequency of data collection is very high in terms of stock volume. Content analysis management plays a vital role in managing the stock data in order to classify and cluster in terms managing the data. The process of data classification and clustering will keep track on the stock in order to satisfy the customer need on demand. The inventory management with respect to supply chain management involves not only controlling the raw materials of stock as well the cost which is related to the stock in the supply chain environment. This process involves in verifying the demand on stock by making use of the concept first in first out (FIFO), Last in First out (LIFO), service in random order (SIRO) or priority treatment techniques in order to verify the demand basis of end user which helps to control the wastages in stock in inventory Management. The error rate and complexity of huge volume of data is very high. We need some techniques in order to prevent the issues which are directly related to the volume and variety of data in managing the stock information within an organization. In this approach, supply chain management and inventory data management deals the huge assortment of data in terms of both volume and variety using different dimensions.

1. Data Classification
2. Data clustering
3. Content analysis
4. Customer retention
5. Inventory based on LIFO, FIFO, SIRO or Priority treatment.

Supply chain Management and inventory data management using big data analytics In inventory management, we support to marketing analysis which helps in identifying the stock with in demand with respect to the end user with the change in need. Based on this survey, we can update the stock management with respect to the time and situation of the end user. Analysis of data prediction is based on customer retention which directly related to the end user satisfaction rate. The increase in data results not only in storage but also in analysing and processing the flow of information in while classifying and clustering the data as per need. There we come up with a concept of content analysis and management which is major aspect in managing the stock within an organization with raised change in demand.

Chapter 4

4. CASE STUDY UNDERTAKEN:

4.1 PROBLEM DEFINITION

Due to the different terminologies used in the literature survey, which is a result of the enormous and rapid growth and development of information and communication technology, we consider it necessary to specify and clarify the definitions used in this article. A computer based transaction processing system is not decision oriented, but it provides two types of reports: a control report that provides information on the errors detected during processing, and monitoring reports on various activities. An information system is a set of organized procedures that, when executed, provide information for decision making and/or control of the organization. It is a combination of hardware, software, people, procedures and data. A computer based information system provides the managers with data processing capabilities and with the information they need to make better decision. An information system contains information and not just data. Information is provided in a meaningful form at the right time. The main impact has been on structured tasks. The value of the information system is the improvement in the value of decision making that can be attributed to using the information system. The term management information system is used interchangeably with IS. An inventory system provides the organizational structure and the operating policies for maintaining and controlling goods to be inventoried. A major drawback for effective inventory management is the lack of accurate and timely information, especially at the operational level. A materials information system is usually designed to reduce lead times, minimize inventory holding and stock- outs, improve financial control, and quickly reschedule material control activities when there are changes in the production plan. The materials information system enables management to use the inventory classification scheme for tighter control, to determine material requirements, to handle schedule changes automatically, to evaluate vendors, and to utilize foreign exchange more effectively. A decision support system is a computer based system that possesses some decision-making or decision-aiding capability. A decision support system provides the decision maker with timely and relevant information. It is an interactive system that provides the user with easy access to decision models and data in order to support semi structured and unstructured decision-making tasks. Decision support systems rely on an integrated set of user friendly hardware and software tools to produce information that is aimed at supporting management in the decision-making process. A decision support system is less structured and provides managers with the flexibility to look into the future and ask “what if” questions. The basic functions of a decision support system for inventory management according to Cook and Russel are the following:

(a) Inventory accounting, which requires a data base and transaction processing system.

(b) Demand forecasting. How much to order and when?

(c) Inventory reporting, Production of reports to be used by inventory analysts, summary reports to higher authority management and ad hoc reports to analyse problem areas.

PROPOSED ALGORITHM

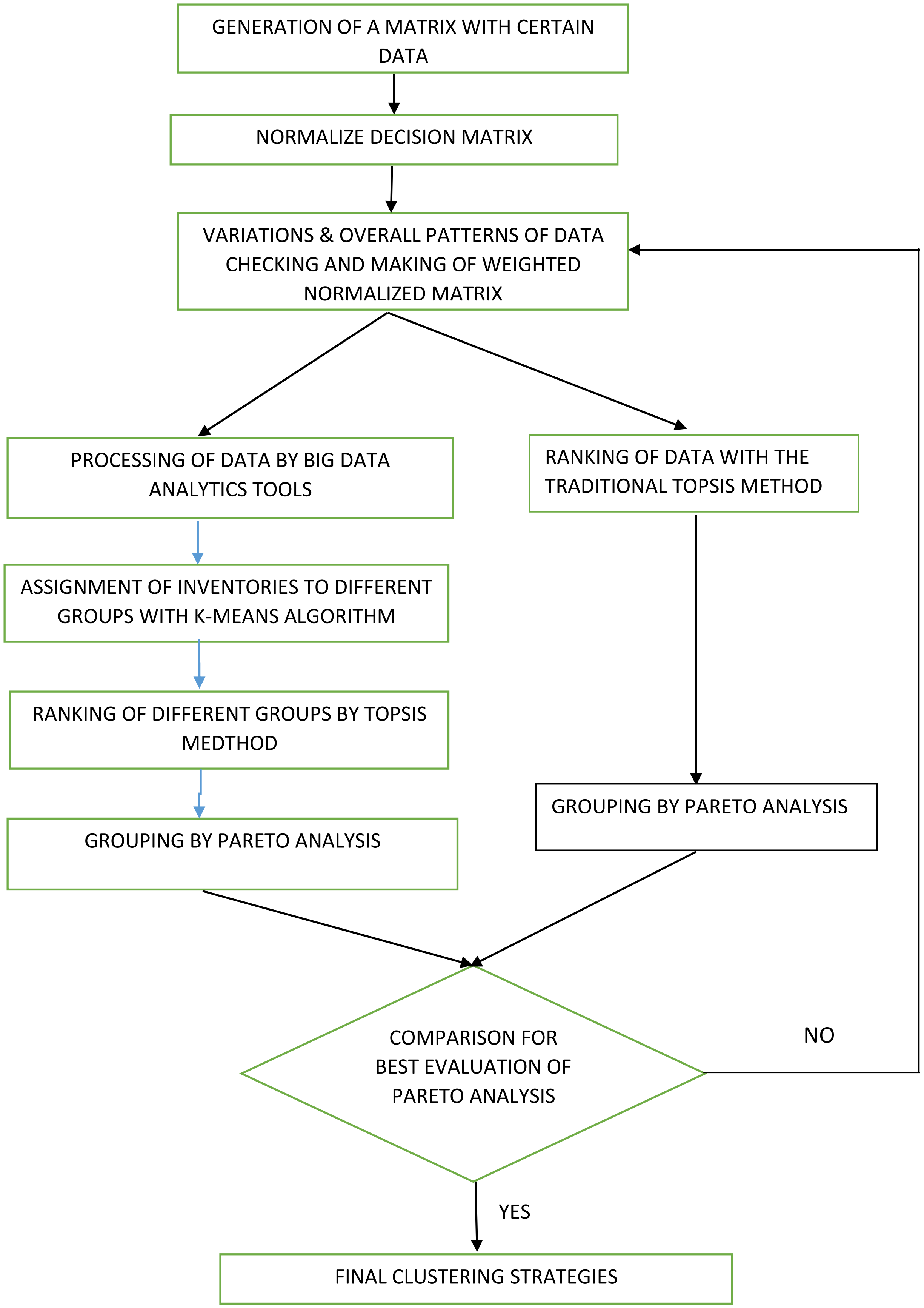


Fig12: Flow chart of the algorithm

An MCIC problem from the literature has been taken and used. This MCIC problem for 47 items was first presented by Flores (1992). Later on, several researchers developed various models to classify inventory based on four criteria: namely, Average Unit Cost (AUC) (-), Average Annual Usage (AAU) (+), Critical Factor (CF) (-) and Lead Time (LT) (-).

Data table:

Table 5: List of items respect to their criteria

Items	AUC	AAU	CF	LT
S1	49.92	5,840.64	1	2
S2	210	5,670.00	1	5
S3	23.76	5,037.12	1	4
S4	27.73	4,769.56	0.01	1
S5	57.98	3,478.80	0.5	3
S6	31.24	2,936.67	0.5	3
S7	28.2	2,820.00	0.5	3
S8	55	2,640.00	0.01	4
S9	73.44	2,423.52	1	6
S10	160.5	2,407.50	0.5	4
S11	5.12	1,075.20	1	2
S12	20.87	1043.50	0.5	5
S13	86.5	1,038.00	1	7
S14	110.4	883.2	0.5	5
S15	71.2	854.4	1	3
S16	45	810	0.5	3
S17	14.66	703.68	0.5	4

Items	AUC	AAU	CF	LT
S18	49.5	594	0.5	6
S19	47.5	570	0.5	5
S20	58.45	467.6	0.5	4
S21	24.4	463.6	1	4
S22	65	455	0.5	4
S23	86.5	432.5	1	4
S24	33.2	398.4	1	3
S25	37.05	370.5	0.01	1
S26	33.84	338.4	0.01	3
S27	84.03	336.12	0.01	1
S28	78.4	313.6	0.01	6
S29	134.34	268.68	0.01	7
S30	56	224	0.01	1
S31	72	216	0.5	5
S32	53.02	212.08	1	2
S33	49.48	197.92	0.01	5
S34	7.07	190.89	0.01	7
S35	60.6	181.8	0.01	3
S36	40.82	163.28	1	3
S37	30	150	0.01	5
S38	67.4	134.8	0.5	3
S39	59.6	119.2	0.01	5
S40	51.68	103.36	0.01	6

Items	AUC	AAU	CF	LT
S41	19.8	79.2	0.01	2
S42	37.7	75.4	0.01	2
S43	29.89	59.78	0.01	5
S44	48.3	48.3	0.01	3
S45	34.4	34.4	0.01	7
S46	28.8	28.8	0.01	3
S47	8.46	25.38	0.01	5

In order to create weighted decision matrix, weighted values of evaluation criteria has been calculated first. The total weighted value of the criteria is equal to 1.

Flores suggested the weight as- AUC= 0.079, AAU= 0.091, CF= 0.420, LT= 0.410

Table 5: weighted normalized matrix

Objective Functions	AUC (-)	AAU (+)	CF (-)	LT (-)
S1	0.00866	0.04073	0.1066	0.028
S2	0.03646	0.03953	0.1066	0.0702
S3	0.00412	0.03512	0.1066	0.0562
S4	0.00481	0.03326	0.00105	0.014
S5	0.01006	0.02426	0.0532	0.0421
S6	0.00541	0.02047	0.0532	0.0421
S7	0.00489	0.01966	0.0532	0.0421
S8	0.00954	0.0184	0.00105	0.0562
S9	0.01275	0.01689	0.1066	0.0843
S10	0.02786	0.01678	0.0532	0.0562
S11	0.00088	0.00749	0.1066	0.028
S12	0.00361	0.00727	0.0532	0.0702
S13	0.01501	0.00723	0.1066	0.0984

Objective Functions	AUC (-)	AAU (+)	CF (-)	LT (-)
S14	0.01916	0.00615	0.0532	0.0702
S15	0.01236	0.00595	0.1066	0.0421
S16	0.00781	0.00564	0.0532	0.0421
S17	0.00254	0.0049	0.0532	0.0562
S18	0.00859	0.00414	0.0532	0.0843
S19	0.00824	0.00396	0.0532	0.0702
S20	0.01014	0.00325	0.0532	0.0562
S21	0.00423	0.00323	0.1066	0.0562
S22	0.01128	0.00316	0.0532	0.0562
S23	0.01501	0.00301	0.1066	0.0562
S24	0.00575	0.00277	0.1066	0.0421
S25	0.00643	0.00257	0.00105	0.014
S26	0.00586	0.00235	0.00105	0.0421
S27	0.01459	0.00233	0.00105	0.014
S28	0.01361	0.00218	0.00105	0.0843
S29	0.02332	0.00187	0.00105	0.0984
S30	0.00971	0.00155	0.00105	0.014
S31	0.01249	0.0015	0.0532	0.0702
S32	0.0092	0.00147	0.1066	0.028
S33	0.00858	0.00137	0.00105	0.0702
S34	0.00122	0.00132	0.00105	0.0984
S35	0.01052	0.00126	0.00105	0.0421
S36	0.00708	0.00113	0.1066	0.0421
S37	0.0052	0.00103	0.00105	0.0702
S38	0.01169	0.00093	0.0532	0.0421
S39	0.01034	0.00082	0.00105	0.0702
S40	0.00896	0.00071	0.00105	0.0843
S41	0.00343	0.00054	0.00105	0.028
S42	0.00654	0.00051	0.00105	0.028

Objective Functions	AUC (-)	AAU (+)	CF (-)	LT (-)
S43	0.00519	0.0004	0.00105	0.0702
S44	0.00838	0.00033	0.00105	0.0421
S45	0.00597	0.00023	0.00105	0.0984
S46	0.005	0.0002	0.0010	0.0421
S47	0.00146	0.00017	0.00105	0.0702

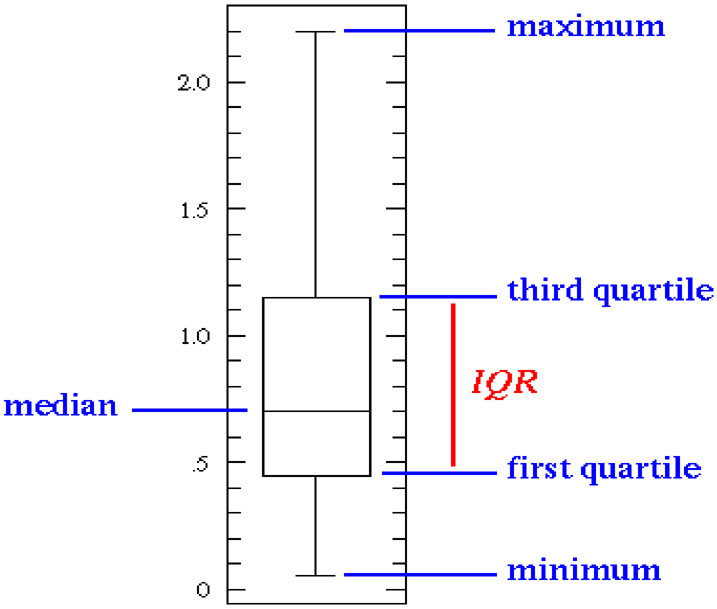
To know the variations & overall patterns of response for a group here box plot graph has been depicted. They provide a useful way to visualise the range and other characteristics of responses for a large group.

4.2 BOX PLOT

In descriptive statistics, a box plot is a method for graphically depicting groups of numerical data through their quartiles.

	AUC	**AAU**	**CF**	**LT**
MIN	0.00088	0.00017	0.00105	0.014
Q1	0.005195	0.001195	0.00105	0.0421
MEDIAN	0.00858	0.00277	0.0532	0.0562
Q3	0.011485	0.00725	0.0799	0.0702
MAX	0.03646	0.04073	0.1066	0.0984
BOX 1 - HIDDEN	0.005195	0.001195	0.00105	0.0421
BOX 2-LOWER	0.003385	0.001575	0.05215	0.0141
BOX 3-UPPER	0.002905	0.00448	0.0267	0.014
WHISKER TOP	0.024975	0.03348	0.0267	0.0282
WHISKER BOTTOM	0.004315	0.001025	0	0.0281

This simplest possible box plot displays the full range of variation (from min to max), the likely range of variation (the IQR), and a typical value (the median). Not uncommonly real datasets will display surprisingly high maximums or surprisingly low minimums called outliers.



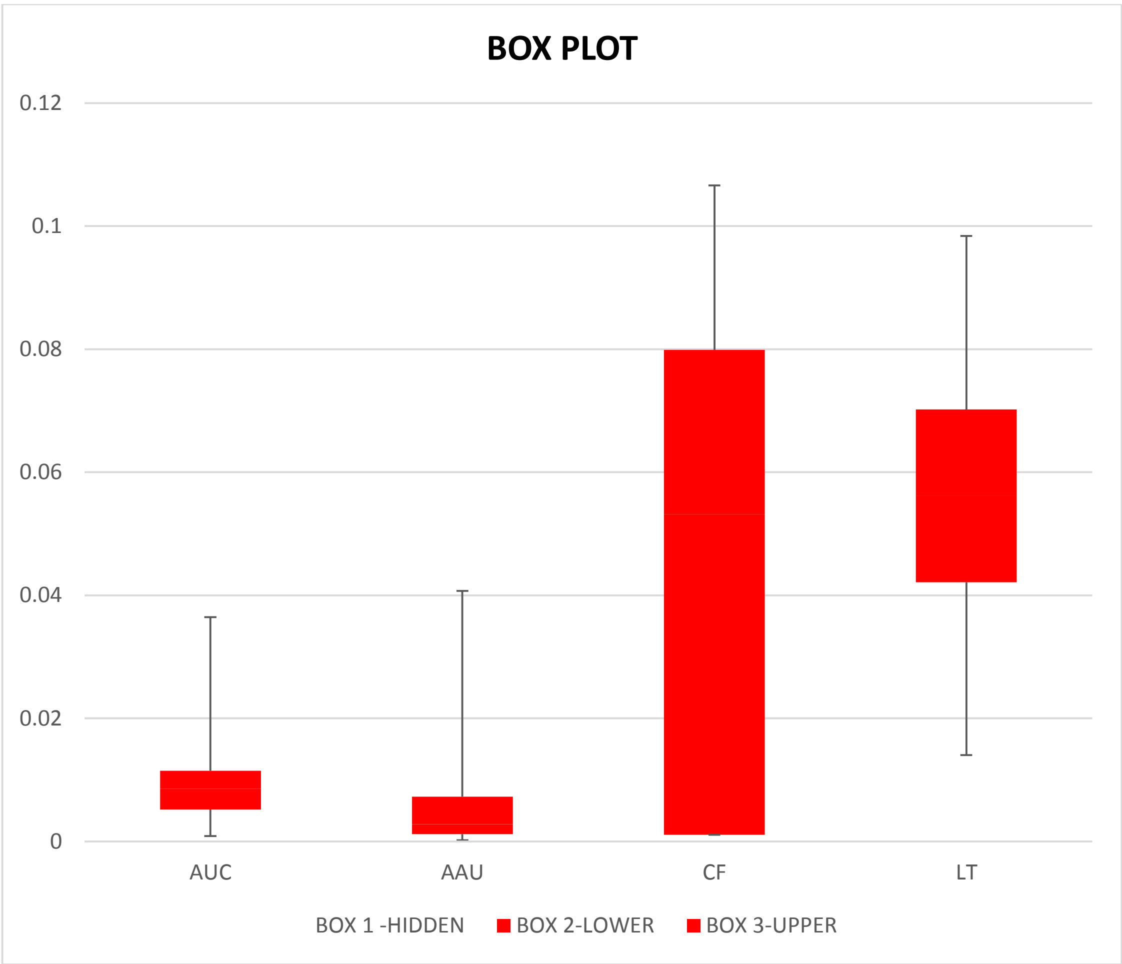


Fig 13: Box plot of the data table

4.3 K-MEANS RESULT:

The K-means clustering in R-software has been used with the help of weighted normalized matrix.

K-means algorithm is an iterative algorithm that tries to partition the dataset into *K* pre-defined distinct non-overlapping subgroups i.e. clusters where each data point belongs to **only one group**. It tries to make the inter-cluster data points as similar as possible while also keeping the clusters as different (far) as possible. It assigns items to a cluster such that the sum of the squared distance between the data points and the cluster's centroid is at the minimum. The less variation we have within clusters, the more homogeneous the data points are within the same cluster.

K-means clustering with 3 clusters of sizes 23, 7 and 17

4.4 Cluster Means:

TABLE 7: CLUSTER MEAN RESULTS OBTAINED FROM R-RESULTS

CLUSTER	AUC	AAU	CF	LT
1	-0.250033763	-0.3234859	0.1268046	-0.6742060
2	0.8233320626	1.9951868	0.7528977	-0.2044061
3	-0.000738111	-0.3838902	-0.4815758	0.9963283

Within cluster sum of squares by cluster:

[1] 42.51415 37.41993 29.20242

(between_SS / total_SS= 40.7 %)

4.5 CLUSTERING VECTOR:

Clustering vector has revealed which class of items belongs to which cluster.

Table 8: Items in their specified clustering vector

Items	Clustering Vector	Items	Clustering Vector	Items	Clustering Vector
S1	2	S2	2	S3	2
S4	2	S5	2	S6	1
S7	1	S8	3	S9	2
S10	2	S11	1	S12	3
S13	3	S14	3	S15	1
S16	1	S17	1	S18	3

Items	Clustering Vector	Items	Clustering Vector	Items	Clustering Vector
S19	3	S20	1	S21	1
S22	1	S23	1	S24	1
S25	1	S26	1	S27	1
S28	3	S29	3	S30	1
S31	3	S32	1	S33	3
S34	3	S35	1	S36	1
S37	3	S38	1	S39	3
S40	3	S41	1	S42	1
S43	3	S44	1	S45	3
S46	1	S47	3		

4.6 TOPSIS of Cluster Means:

In the columns of each criteria in the weighted normalized decision matrix, maximum values are determined as ideal (+ve) solution values and minimum values are determined as negative ideal solution (-ve).

TABLE 9: Formulation of cluster means for TOPSIS method

	AUC (-)	AAU (+)	CF (-)	LT (-)
CLUSTER 1	-0.250033763	-0.3234859	0.1268046	-0.6742060
CLUSTER 2	0.8233320626	1.9951868	0.7528977	-0.2044061
CLUSTER 3	-0.000738111	-0.3838902	-0.4815758	0.9963283
PIS	-0.250033763	1.9951868	-0.4815758	-0.6742060
NIS	0.8233320626	-0.3838902	0.7528977	0.9963283

Table 10: Determination of Separation Measure (SM)

	S_i^+	S_i^-
CLUSTER 1	5.33265	4.33479
CLUSTER 2	2.896750	7.1017704
CLUSTER 3	8.512840	2.2030164

$$SM_i = \frac{S_i^-}{S_i^+ + S_i^-}; i = 1,2, \dots, N$$

Table 11: Findings of Relative closeness index

CLUSTER 1	0.448390
CLUSTER 2	0.710275
CLUSTER 3	0.2055847

DECISION BASED ON THE RESULTS:

Table 12: ABC categorization of clusters

CLUSTER	TOTAL ITEMS IN CLUSTER	CATEGORY
CLUSTER 2	7	A
CLUSTER 1	23	B
CLUSTER 3	17	C

VISUALIZATION OF CLUSTEING COMPONENTS IN CIRCULAR SHAPES:

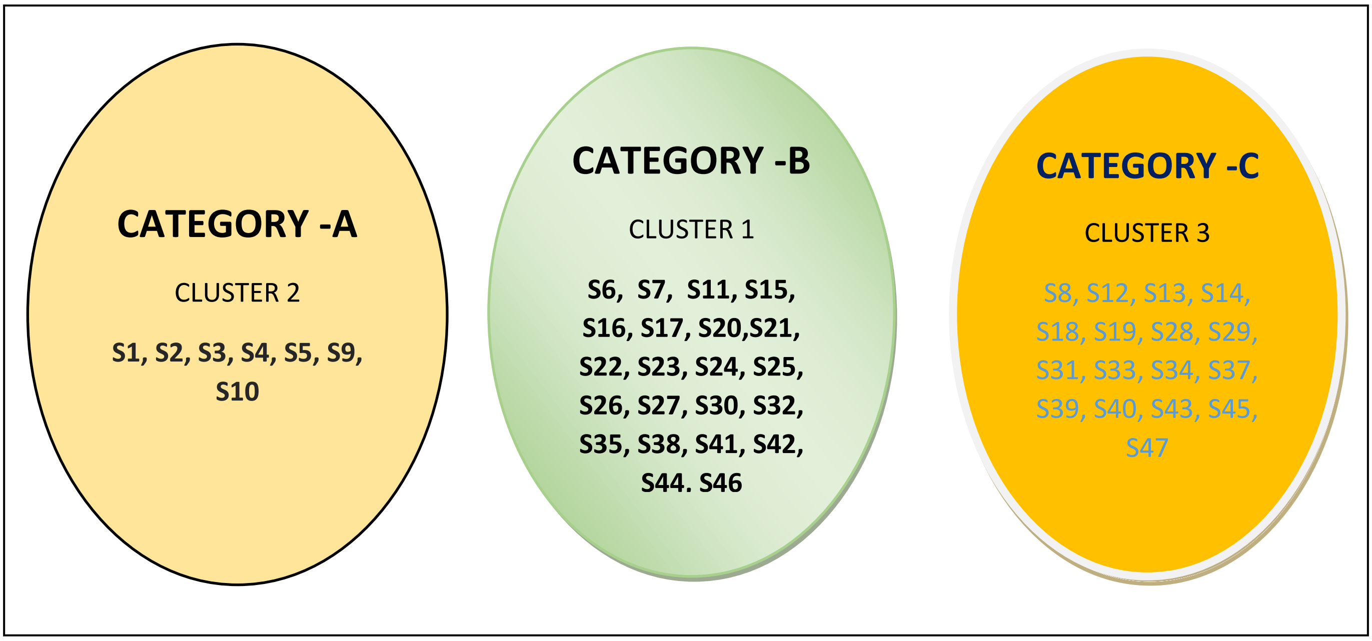


Table 13: Comparison of Traditional TOPSIS & TOPSIS with Big Data Analytics with augmentation of ABC analysis

SL NO	Objective functions	Traditional TOPSIS with ABC classification	Objective functions	TOPSIS with Big Data Analytics & ABC classification
1	S1	A	S1	A
2	S3	A	S3	A
3	S9	A	S9	A
4	S2	B	S2	A
5	S4	C	S4	A
6	S5	B	S5	A
7	S10	B	S10	A
8	S18	C	S17	B
9	S12	B	S20	B
10	S19	B	S22	B
11	S31	B	S30	B
12	S13	B	S27	B
13	S14	C	S25	B
14	S47	C	S42	B
15	S43	C	S41	B
16	S37	C	S46	B
17	S8	C	S44	B
18	S33	C	S26	B
19	S39	C	S35	B
20	S40	C	S11	B
21	S34	C	S32	B
22	S45	C	S24	B
23	S28	C	S36	B
24	S29	C	S15	B
25	S17	B	S21	B
26	S20	B	S23	B
27	S22	B	S38	B
28	S30	C	S7	B
29	S27	C	S6	B
30	S25	C	S16	B

SL NO	Objective functions	Traditional TOPSIS with ABC classification	Objective functions	TOPSIS with Big Data Analytics & ABC classification
31	S42	C	S18	C
32	S41	C	S12	C
33	S46	C	S19	C
34	S44	C	S31	C
35	S26	C	S13	C
36	S35	C	S14	C
37	S11	A	S47	C
38	S32	A	S43	C
39	S24	A	S37	C
40	S36	A	S8	C
41	S15	A	S33	C
42	S21	A	S39	C
43	S23	A	S40	C
44	S38	B	S34	C
45	S7	B	S45	C
46	S6	B	S28	C
47	S16	B	S29	C

Chapter 5

5. Discussion:

5.1 Cluster Dendrogram

I am using the default method of hclust, which is to update the distance matrix using what R calls "complete" linkage. Using this method, when a cluster is formed, its distance to other objects is computed as the maximum distance between any object in the cluster and the other object. Other linkage methods will provide different solutions, and should not be ignored. For example, using method=ward tends to produce clusters of fairly equal size, and can be useful when other methods find clusters that contain just a few observations.

Now that we've got a cluster solution (actually a collection of cluster solutions), how can we examine the results?

The main graphical tool for looking at a hierarchical cluster solution is known Dendrogram. This is a tree-like display that lists the objects which are clustered along the x-axis, and the distance at which the cluster was formed along the y-axis. (Distances along the x-axis are meaningless in a Dendrogram; the observations are equally spaced to make the Dendrogram easier to read. To create a Dendrogram from a cluster solution, simply pass it to the plot function. The result of this research data table is displayed below.

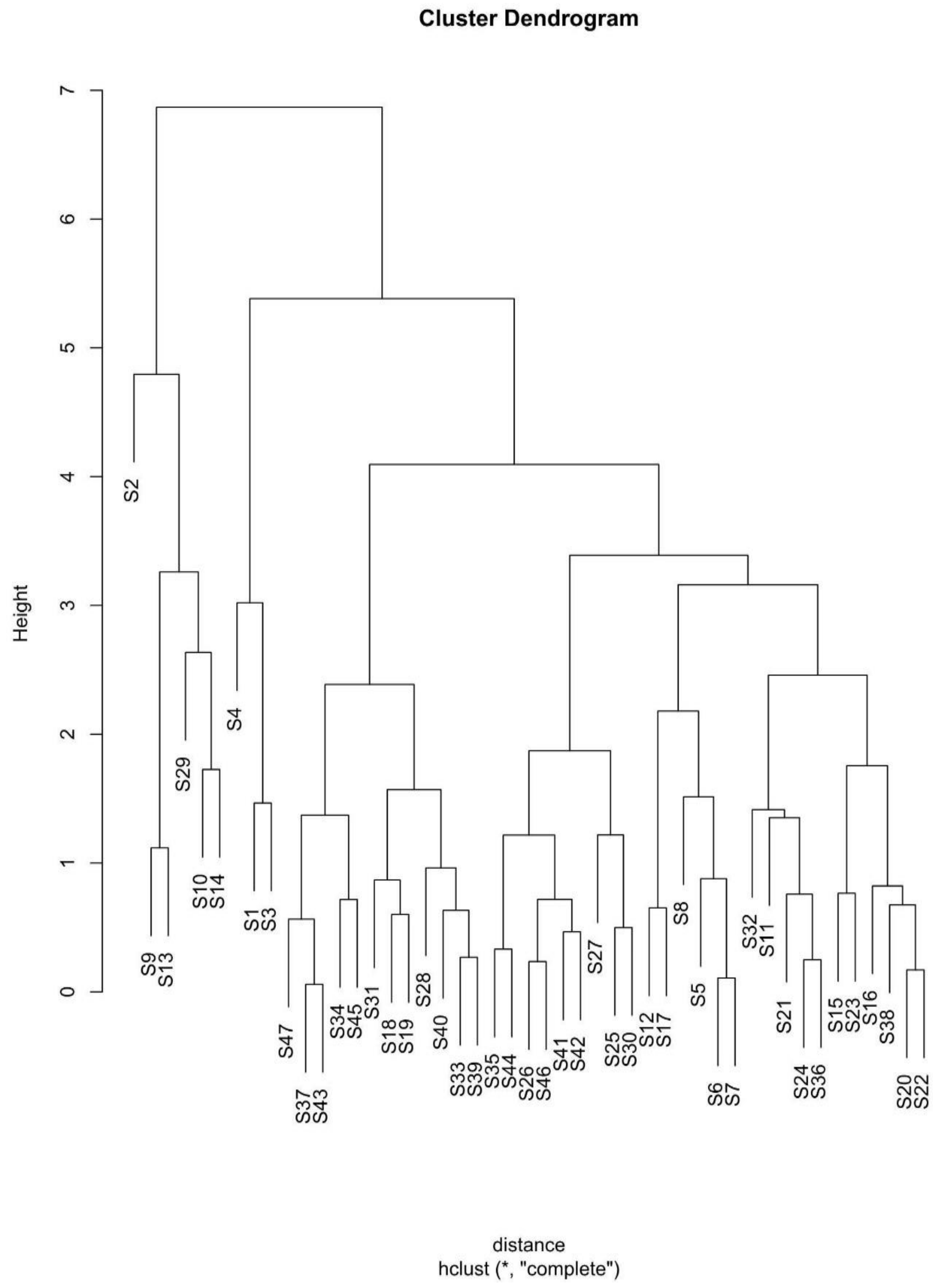


Fig. 14: cluster Dendrogram based on complete linkage method

If you choose any height along the y-axis of the Dendrogram, and move across the Dendrogram counting the number of lines that you cross, each line represents a group that was identified when objects were joined together into clusters. That defines a three-cluster solution; by following the line down through all its branches, we can see the names of the inventory that are included in these three clusters. Since the y-axis represents how close together observations were when they were merged into clusters, clusters whose branches are very close together (in terms of the heights at which they were merged) probably aren't very reliable. But if there's a big difference along the y-axis between the last merged clusters and the currently merged one, which indicates that the clusters formed are probably doing a good job in showing us the structure of the data. Looking at the Dendrogram for the inventory data, there are clearly three very distinct groups; the right hand groups seems to consist of two more distinct cluster, while most of the observations in the left hand group are clustering together at about the same height For this data set, it looks like either three groups might be an interesting place to start investigating. This is not to imply that looking at solutions with more clusters would be meaningless, but the data seems to suggest that three clusters might be a good start. For a problem of this size, we can see the number of inventory items, so we could start interpreting the results immediately from the Dendrogram, but when there are larger numbers of observations, this won't be possible.

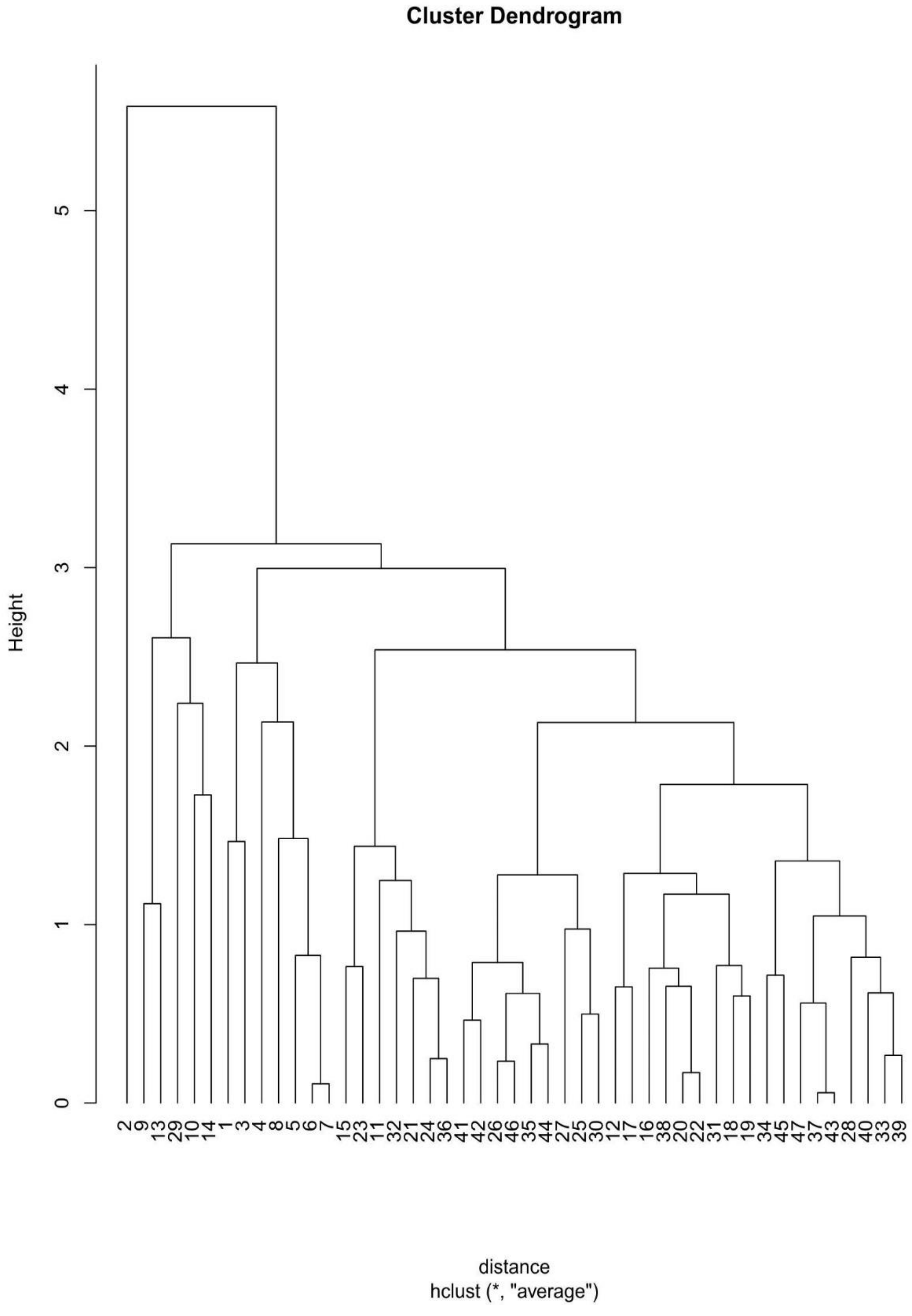


Fig. 15: cluster Dendrogram based on average linkage method

5.2 SILHOUETTE PLOT

Silhouette refers to a method of interpretation and validation of consistency within clusters of data. The technique provides a succinct graphical representation of how well each object has been classified. The silhouette value is a measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation). The silhouette ranges from −1 to +1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighbouring clusters. If most objects have a high value, then the clustering configuration is appropriate. If many points have a low or negative value, then the clustering configuration may have too many or too few clusters.

The silhouette can be calculated with any distance metric, such as the Euclidean distance or the Manhattan distance.

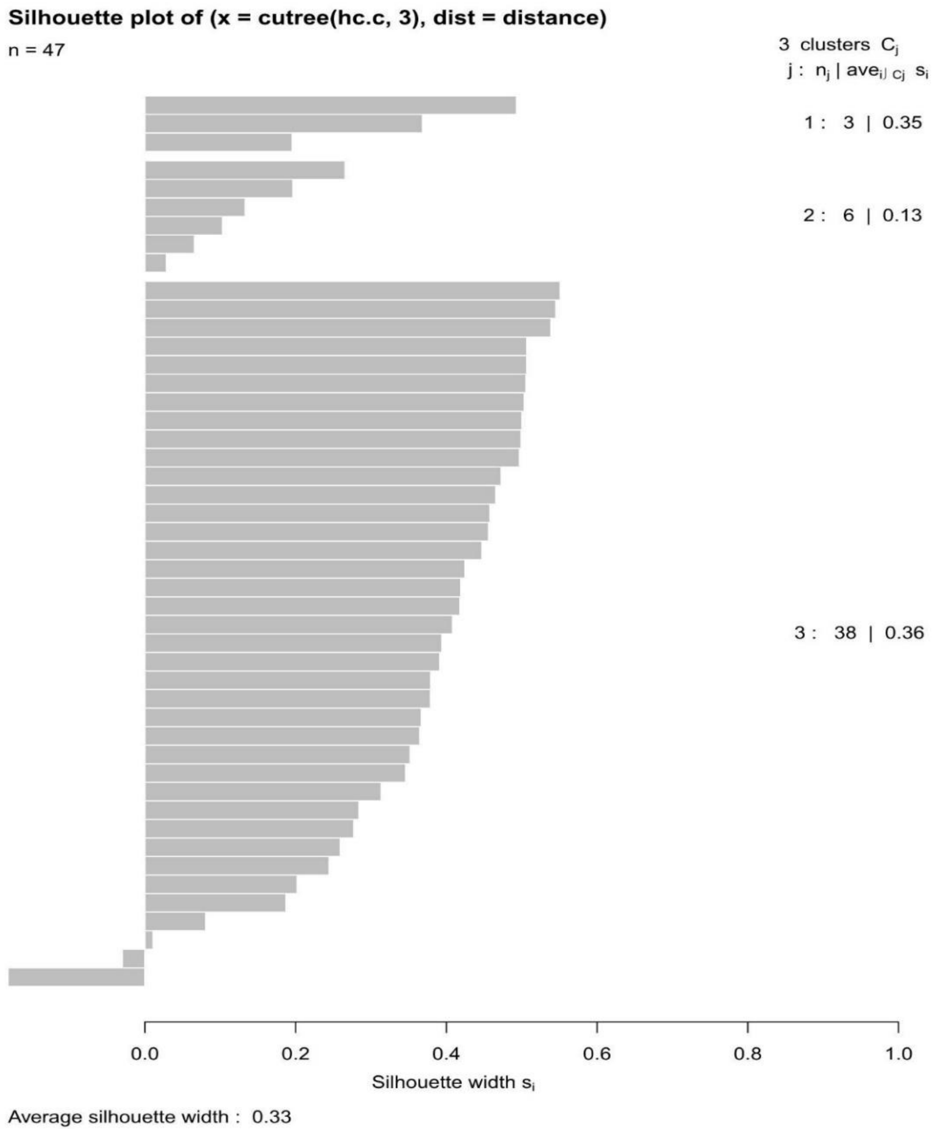


Fig. 16: silhouette plot of three clusters

Table 12: A Comparison of ABC classification using TOPSIS, Annual Dollar usage and AHP weighted score

Items	TOPSIS with K-Means clustering	ABC Classic	AHP Weighted
S1	A	A	A
S3	A	A	A
S9	A	A	A
S2	A	A	A
S4	A	A	C
S5	A	A	B
S10	A	A	B
S17	B	B	B
S20	B	B	B
S22	B	B	B
S30	B	C	C
S27	B	C	C
S25	B	C	C
S42	B	C	C
S41	B	C	C
S46	B	C	C
S44	B	C	C
S26	B	C	C
S35	B	C	C
S11	B	B	B
S32	B	C	B
S24	B	B	A
S36	B	C	B
S15	B	B	A
S21	B	B	A
S23	B	B	A
S38	B	C	C
S7	B	A	C
S6	B	A	C
S16	B	B	B
S18	C	B	A
S12	C	B	B
S19	C	B	B
S31	C	C	B
S13	C	B	A

Items	TOPSIS with K-Means clustering	ABC Classic	AHP Weighted
S14	C	B	B
S47	C	C	C
S43	C	C	C
S37	C	C	C
S8	C	A	C
S33	C	C	C
S39	C	C	C
S40	C	C	C
S34	C	C	C
S45	C	C	B
S28	C	C	C
S29	C	C	C

5.3 SCOR model

This research provides a new description of the inventory management model, incorporating big data and inventory management. The SCOR model, as developed by the Supply Chain Council, portrays four essential pillars, i.e. planning, sourcing, making and delivering, which involve the flow of finance, material movement and information flow to integrate demand and supply management across the supply chain. This integration requires proper collaboration and strategic thinking, which involve planning, target setting, monitoring and control. For optimal utilization of the system, in process inventory needs to be eliminated, which will help to reduce costs and generate supply chain surplus. This signals access to relevant and timely information and ensures efficient data management for the regulation of activities and performance. The SCOR model is suitable for the evaluation of the financial performance of inventory management and can provide a practical decision support tool for environmental assessment and examining competing decision alternatives along the chain.

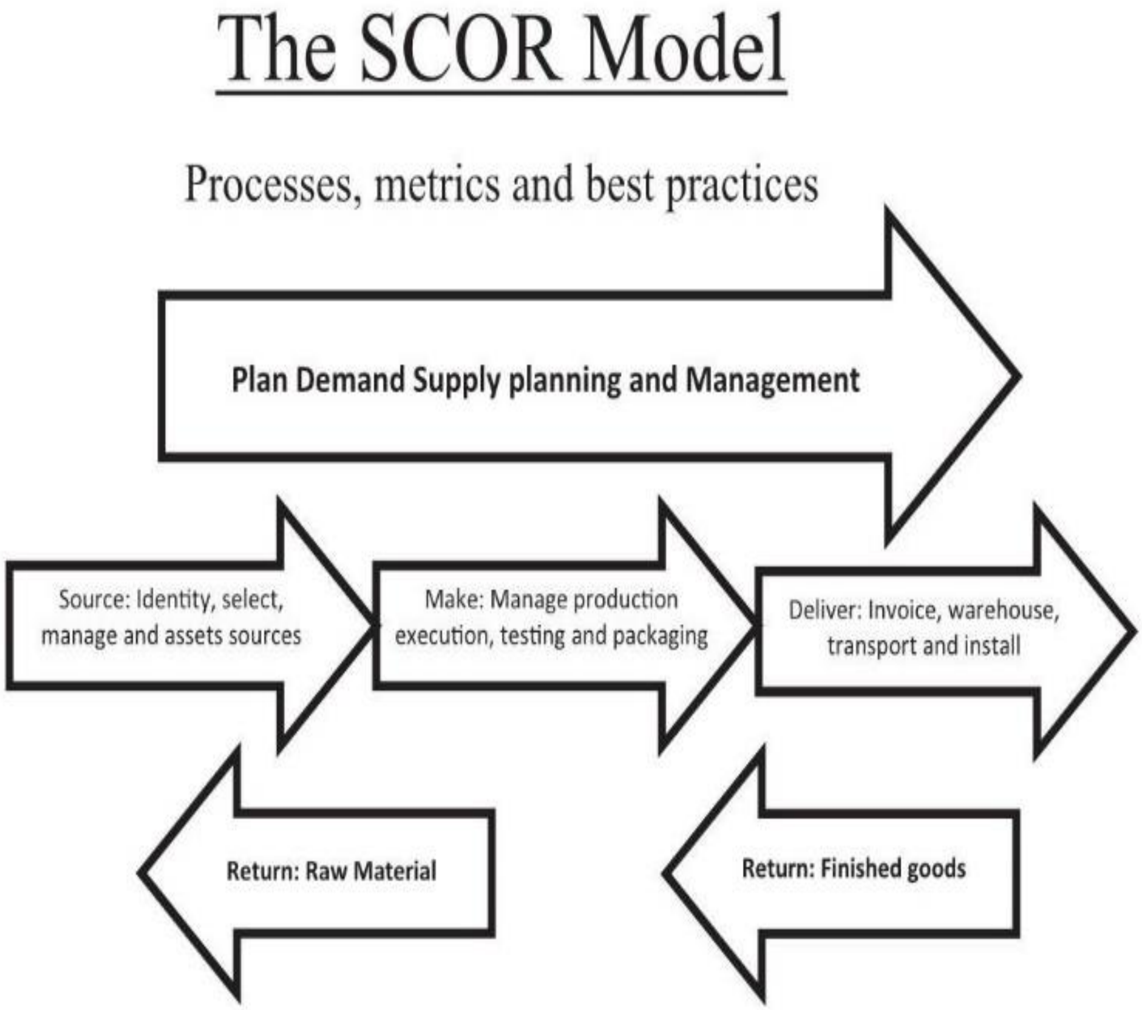


Figure 17: The SCOR model (Supply Chain Council: SCOR 9.0 Overview Booklet, 2008)

Plan

- P1: Use big data tools to assess customer demand and needs in real time.
- P2: Use“4V” framework consisting of high volume, high velocity, high veracity and high variety information, which uses various processing measures to ensure better decision making.
- P3: Use analytics to improve processes and ensure business optimization.

Source

- S1: Gather tremendous amounts of data from various processes, i.e., the use of sensors, RFID and tracking devices, which serves as a breeding ground for the generation of big data.
- S2: Use media to rate various vendors.
- S3: Use analytic hierarchy processing (AHP) to determine the varying degree of importance of the flow of inputs and provides a strategic perspective in rating suppliers.

Make

- ✓ M1: Use big data to strengthen market competitiveness and improve data quality management and data usage experience.
- ✓ M2: Use big data approach to decision making for logistics planning and scheduling using data collected via the use of RFID or sensors on manufacturing shop floors.

Deliver

- ❖ D1: Use big data to improve visibility by providing an integrated framework to monitor performance and customer interaction through real time data analysis and critical decision-making scenarios, thus mitigating risk and inventory management Stockout and failures

Return

- R1: Exploit the huge amount of collected data with the help of information technology (IT), i.e., business intelligence insights, analytics, etc., improve existing supply chain practices, reduce costs and provide better inventory management, in turn increase profits in the supply chain industry.
- R2: Use demand management to deal with the obsolescence of goods and wastage, especially for perishable goods, for which both agile and lean methodologies may be used.

Chapter 6

6.1 CONCLUSION:

The traditional paradigm for inventory management is to develop R-TOPSIS to generate forecasts that accurately predict the value and the level of uncertainty of future demand. These forecasts are then used as an input to an optimization problem that evaluates trade-offs and respects constraints in order to come up with decisions about managing materials. This two-step process, which is embodied in all current material-management planning and control systems, can be replaced by a single-step process that looks for the best relationship among all of the data and the decisions. Based on learning from the past, a "best" relationship can be identified, which will generate decisions, as future uncertainty is resolved, that are better than the decisions derived from the traditional two-step approach of first forecast and then optimize.

The classical single criterion ABC inventory classification is simple, straightforward. However, in the management of an inventory, other criteria influence its administration. Factors such as lead time, criticality and obsolescence may sometimes be so important that they override financial concerns. When considering multiple criteria, a major concern is not to complicate the managerial process. However, it is necessary to incorporate multiple criteria and still have a reasonably simple set of inventory policies. The utilization of R-TOPSIS provides a way to combine these multiple criteria. In this paper, a simple approach based on MCDM technique for inventory classification is proposed when multiple criteria are considered. The method is easy to understand by inventory managers and very effective while considering multiple criteria. An inventory item measured under different parameters ranks differently from parameter to parameter. An item may be very costly, but not as critical. Another may be very critical but have a very short lead time. When all of these factors are synthesized into a single ranking, very likely the assignment of items to classes will be more uniform. One of the limitations of this method is an assignment of weights to the criteria. For some cases assignment of weights can be difficult and takes more time. However, use of multiple criteria ABC analysis can improve the quality and completeness of the inventory analysis.

6.2 REFERENCES:

[1] Anand Rabi Dhar, Bijan Sarka**.**Multi-Criteria ABC Inventory Classification Using MULTIMOORA and Extended MULTIMOORA Considering Fuzzy Expert Opinions.

[2] Bhattacharya, A., Sarkar, B., & Mukherjee, S. K. (2007). Distance-based consensus method for ABC analysis. International Journal of Production Research, 45(15), 3405-3420.

[3] Blaz Zupan, Janez Demsar (2008)Open-Source Tools for Data Mining; clinicsin laboratory medicines, Clin Lab Med 28 (2008) 37–54.

[4] Braglia, M., Grassy, A., & Montana, R. (2004). Multi-attribute classification method for spare parts inventory management. Journal of Quality in Maintenance Engineering, 10(1), 55-65.

[5] Brauers, W.K.M., Baležentis, A., Baležentis, T. (2011). MULTIMOORA for the EU member states updated with fuzzy number theory. Technology and Economic Development, 17, 259–290.

[6] Brauers, W.K.M., Kildienė, S., Zavadskas, E.K., & Kaklauskas, A. (2013). The construction sector in twenty European countries during the recession 2008–2009–country ranking by MULTIMOORA. International Journal of Strategic Property Management, 17, 58–78.

[7] Brauers, W.K.M., & Zavadskas, E.K. (2006). The MOORA method and its application to privatization in a transition economy. Control and Cybernetic Journal, 35, 445–469.

[8] Christopher, M., and L. J. Ryals. 2014. "The Supply Chain Becomes the Demand Chain." Journal of Business Logistics 35 (1): 29–35.

[9] Chui, M., Kamalnath, V., & McCarthy, B. (2018). An executive's guide to AI. McKinsey & Company. Available at http://www. mckinsey.com/business-functions/mckinsey-analytics/ our-insights/an-executives-guide-to-ai.

[10] CAI, J., X. Liu, Z. Xiao, and J. Liu. 2009. "Improving Supply Chain Performance Management: A Systematic Approach to Analysing Iterative KPI Accomplishment." Decision Support Systems 46 (2): 512–521.

[11] Canetta, L., Cheikhrouhou, N., & Glardon, R., (2005). Applying two-stage SOM-based clustering approaches to industrial data analysis, Production Planning & Control, 16(8), 774-784.

[12] Cook, T.M. and Russell, R.A., 1989. Introduction to Management Science. Prentice-Hall, Englewood Cliffs, NJ.

[13] Dimitris Bertsimas, Nathan Kallus, Amjad Hussain (2016). Inventory Management in the Era of Big Data. Production and Operations Management 0(0), pp. 1–12, © 2016 Production and Operations Management Society.

[14] Ernst, R., & Cohen A., (1990) Operations related groups (ORGs): a clustering procedure for production/ inventory systems. Journal of Operations Management (4), 574-598.

[15] Flores, B.E., & Whybark, D.C. (1986). Multiple Criteria ABC Analysis, International Journal of Operations and Production Management, 6(3), 38-46.

[16] Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information and Management, 53(8), 1049—1064.

[17] Gajpal, P. P., Ganesh, L. S., & Rajendran, C. (1994). Criticality analysis of spare parts using the analytic hierarchy process. International Journal of Production Economics, 35(1), 293-297.

[18] Hoole, R. 2005. “Five Ways to Simplify Your Supply Chain.” Supply Chain Management: An International Journal 10 (1): 3–6.

[19] Hadi-Vencheh, A., & Mohamadghasemi, A. (2011). A AHP-DEA approach for multiple criteria ABC inventory classification. Expert .system with applications, 38, 3346-3352.

[20] Huiskonen, J., Niemi, P., & Pirttila, T. (2005). The role of C-products in providing customer service -refining the inventory policy according to customer-specific factors. International Journal of Production Economics, 93-94(1), 139-149.

[21] Kietzmann, J., Paschen, J., & Treen, E. (2018). Artificial intelligence in advertising: How marketers can leverage artificial intelligence along the consumer journey. Journal of Advertising Research, 58(3), 263—267.

[22] Lingaraj, B.P. and Balasubramanian, R., 1983. An inventory management and materials information system for aircraft production. Interfaces 13(5): 65570.

[23] Li, H., & Musings, S. (2017, April 12). Which machine learning algorithm should I use? SAS. Available at http://blogs.sas.com/content/subconsciousmusings/2017/04/12/ machine-learning-algorithm-use.

[24] Manthou, V. (1994). Concepts and applications of inventory management in Northern Greece. International Journal of Production Economics, 35(1-3), 149–152. Doi: 10.1016/0925-5273(94)90075-2.

[25] Mudassar RAUF, Zailin GUAN, Shoaib SARFRAZ, Jabir MUMTAZ, Sulaiman ALMAIMAN, Essam SHEHAB and Mirza JAHANZAIB(2018) MCIC Based on Multi-Criteria Decision-Making (MCDM) Technique. Advances in Manufacturing Technology XXXII P. Thorvald and K. Case (Eds.) IOS Press, 2018.

[26] Ntabe, E. N., L. LeBel, A. D. Munson, and L. A. Santa-Eulalia. 2015. “A Systematic Literature Review of the Supply Chain Operations Reference (SCOR) Model Application with Special Attention to Environmental Issues.” International Journal of Production Economics 169: 310–332.

[27] Ng, W. L. (2007) A Simple classifier for multiple criteria ABC analysis European Journal of Operational Research, 177(1), 344-353.

[28] Pooya Tabesh, Elham Mousavidin, Sona Hasani (2018). Implementing big data strategies: A managerial perspective. BUSHOR-1561; No. of Pages 12, www.elsevier.com/locate/bushor.

[29] Partovi, F. Y., & Anandarajan, M. (2002). Classifying inventory using an artificial neural network approach. Computers & Industrial Engineering, 41(4), 389-404.

[30] Partovi, F.Y., & Burton, J.(1993).Using the analytic hierarchy process for ABC analysis. International Journal of Operation & Production management, 13 (9), 29 -44.

[31] Reid, R. A. (1987). The ABC method in hospital inventory management: A practical approach. Production and Inventory Management, 28 (4), 67—70.

[32] Supply Chain Council. 2008. SCOR 9.0 Overview Booklet.

[33] Seetha Raman, Nitin Patwa, Indu Niranjan, Ujjwal Ranjan, Krishna Moorthy & Ami Mehta(2018)Impact of big data on supply chain management by published in International Journal of Logistics Research and Applications, A Leading Journal of Supply Chain Management.

[34] Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. Journal of Business Research, 70, 263—286.

[35] V. Manthou, 1994. Concepts and applications of inventory management in Northern Greece. Int. J. Production Economics 35 (1994) 149-152.

[36] Vasiliki Balioti, Christos Tzimopoulos and Christos Evangelides. Multi-Criteria Decision Making Using TOPSIS Method Under Fuzzy Environment. Application in Spillway Selection. EWaS International Conference on “Insights on the Water-Energy-Food Nexus”, Lefkada Island, Greece, 27–30 June 2018.

[37] Waller, M. A., and S. E. Fawcett. 2013. “Click Here for a Data Scientist: Big Data, Predictive Analytics, and Theory Development in the Era of a Maker Movement Supply Chain.” Journal of Business Logistics 34 (4): 249–252. doi:10.1111/jbl.12024.

[38] Wang, G., A.Gunasekaran, E.W.Ngai, andand, and T.Papadopoulos.2016. “Big Data Analytics in Logistics and Supply Chain Management: Certain Investigations for Research and Applications. “International Journal of Production Economics 176: 98–110.

[39] Zeng, J., & Glaister, K. W. (2017). Value creation from big data: Looking inside the black box. Strategic Organization, 16(2), 105—140.

[40] Zhong, R.Y.,G.Q.Huang,S.Lan,Q.Dai,X.Chen,andT.Zhang.2015.“A Big Data Approach for Logistics Trajectory Discovery from RFID-Enabled Production Data.” International Journal of Production Economics, 165260–165272. doi:10.1016/j.ijpe.2015.02.014.

[41] Zhou, P., & Fan (2007). A not: on multi-criteria ABC inventory classification using weighted linear optimization. European Journal of Operational Research, 182(3), 1488-1491.

6.3 FUTURE SCOPE:

In Inventory cum supply chain management, costs are driven by extreme events: it's the surprisingly high demand that generates stock-outs and customer frustration, and the surprisingly low demand that generates dead inventory and consequently costly inventory write-offs. As all executives know, businesses should hope for the best, but prepare for the worst. When the demand is exactly where it was expected to be, everything goes smoothly. However, the core forecasting business challenge is not to do well on the easy cases, where everything will be going well, even considering a crude moving average. The core challenge is to handle the tough cases; the ones that disrupt your supply chain, and drive everybody nuts.

Back in 2016, Lokad developed a radically new way of forecasting, namely probabilistic forecasts. Then, more recently those forecasts got another massive upgrade through deep learning. Simply put, a probabilistic forecast of demand does not merely give an estimate of the demand, but assesses the probabilities of every single future. The probability of zero units of demand is estimated, the probability of 1 unit of demand is estimated, of 2 units of demand, and so on... Every level of demand gets its estimated probability until the probabilities become so small that they can safely be ignored.

These probabilistic forecasts provide an entirely new way of looking at the future. Instead of being stuck in a wishful thinking perspective, where forecast figures are expected to materialize, probabilistic forecasts remind you that everything is always possible, just not quite equally probable.

The future scope of the Inventory Management System should be the Integration of Accounts Receivable, Accounts Payable, and General Ledger allowing Inventory count while continuing to process transactions with deep learning.

Another research area can be Inventory adjustments by quantities and costs allowing importing of inventory items for rapidly adding new products with big data analytics.