An enhanced approach of Hyper ETL to make Efficient Decision Making in Data Mart using Decision Analysis Criteria.

[1]A. PREMA, [2]A.PETHALAKSHMI

[1]RESEARCH SCHOLAR, DEPARTMENT OF COMPUTER SCIENCE,
RESEARCH AND DEVELOPMENT CENTRE,BHARATHIAR UNIVERSITY
COIMBATORE,TAMILNADU,INDIA,

[2]ASSOCIATER PROFESSOR, MVM GOVERNEMNT ARTS COLLEGE
DINDIGUL,TAMILNADU,INDIA

Abstract

In this paper, a straightforward, painless to use and systematic process, rooted in Hyper ETL (Extraction, Transformation and Load) tool, is put forward to assist in the selection among choices, with more than a few decision criteria. In order to expand significant selling reward, it is necessary to study the application of decision support system based on data mart. Hereby, we discuss the various decision analysis methods to take efficient decision making in data mart which hike revenue in the business. We construct ideas on how to put into action, and demonstrate them with sales records examples. The purpose of a vital sales data mart using Hyper ETL is to offer decision-makers entrée to steady, trustworthy, and apt data for systematic forecast and planning.

Keywords: Hyper ETL, Decision support System, decision making, Data mart.

1. INTRODUCTION

As we survive in an era of scientific development, technical innovation and earlier advancement in modern computing, with great interest utmost care is taken in this paper to help aspiring academic and analysis methodologies in sales data mart. This ETL tool is used to simplify the process of migrating data, standardize the method of data migration, and store all data transformation logic as Meta data. In turn it enable the users, managers and architects to understand, review, and modify the various interfaces and reduce the cost and effort associated with building interfaces. Extraction is the process of reading data from a specified source database and extracting a desired subset of data. Transformation phase applies a chain of rules or functions to the extracted data to derive the data to be loaded. Three forms of transformations are utilized, that is subsets of tables, formatting data and primary keys and indexes. Subsets are created to remove personally individual information. Loading is the process of writing the data into the target database. Data mart is talented as one of the most recent region of growth in large scale industry and the vital mechanism and tool for business oriented information for future analysis and decision-making. A data mart is a persistent physical store of operational and aggregated data statistically processed data that supports business people in making decisions based primarily on analysis of past activities and results. A data mart contains a predefined subset of enterprise data organized for rapid analysis and reporting and collects data on a specific subject area such as sales or production or accounts or Human Resource management or customer information. It can be a subset of company data warehouse and it is proposed to meet the desires of a single department. An operational data store is an updatable set of integrated data used for enterprise-wide strategic decision making. The term ‘data mart’ probably takes advantage over the term ‘data warehouse’. The whole data warehousing process is about creating data bringing that knowledge to people. A warehouse is a place where things are stored away. A mart is a place where something is stored.
Section 2 of this paper deals with related work done in the Extract, Transformation and Loading into the data warehouses. Section 3 explains an actual process of Extract, Transform and Load. Section 4 explains an approach of Hyper ETL. In section 5, Experimental analysis and implementation are given, and finally, section 6 presents a conclusion of this paper.

2.Related Works:

A data warehouse is a storage area of operational data that has been extracted from original electronic sources and transformed so that query, analysis and reporting on trends within historic data are possible and efficient. The analyses provided by data warehouses may involve strategic planning, decision support, and monitoring the outcomes of a chosen strategy. Typically, data that is loaded into a data warehouse is derived from diverse sources of operational data, which may consist of data from databases, feeds, application files or flat files. [2] The data must be extracted from these diverse sources, transformed to a common format, and loaded into the data warehouse. Typically, it is further aggregated into a data mart for efficient reporting. The ETL (Extract, Transform and Load) process is a critical step in any data warehouse implementation, and continues to be an area of major significance whenever the ETL code is updated. Once the data warehouse and data marts are populated, business intelligence applications facilitate querying, analysis and reporting. The business intelligence tools may provide simple presentations of data based on queries, or may support sophisticated statistical analysis options. Data warehouses may have multiple front-end applications, depending on the desires of the user society.

A data warehouse, however, requires a concise, subject-oriented scheme that facilitates online data analysis. The simplest scheme is a single table scheme, which consists of redundant fact tables. The most common modeling paradigm according to this is star schema, in which the data warehouse contains a large central fact table containing the bulk of data, with no redundancy, and a set of smaller attendant tables one for each dimension. Snowflake schema is a variant of star schema model, where some dimension tables are normalized; causing thereby further splitting the data into additional tables [10].

The topic of data warehousing encompasses application tools, architectures, and information service and communication infrastructures to synthesize information useful for decision-making from distributed heterogeneous operational data sources. The information is brought together into a single repository, called a data warehouse (DW), suitable for direct querying and analysis and as a source for building logical data marts oriented to specific areas of the enterprise [15] F.McGuff, proposed an approach to the design of DWs based on a business model of the enterprise which was actually a relational database scheme. Regretfully, conceptual and logical designs are mixed up; since logical design is necessarily targeted towards a logical model, no unifying conceptual modeling issues such as the structure of attribute hierarchies and non additive constraints. The approach to conceptual DW modeling shared several ideas with his early work and neglects other conceptual issues such as additive and scheme overlapping [16].

Milija, Milution, and Milan described about the design and implementation of data warehouse as well as the use of data mining algorithms for the purpose. This system represents a good base for analysis and predictions in the following time period for the purpose of quality business decision-making by top management. In this paper, they dealt about the steps in designing and development of data warehouse and then implementation of data mining algorithms for the purpose of deducting rules, patterns and finally, explained the knowledge as a resource for support in the process of decision making. They conclude the paper that show the phases through which a DW and DM solution is formed and DW offers a flexible solution to the user, who can use tools, like Excel, with user-defined queries to explore the database more efficiently in comparison to all other tools form the OLTP environment. This approach in data analysis becomes more and more popular because it enables OLTP systems to get optimized for their purpose and to transfer data analysis to OLAP systems. [3]

Robert Winter and Markus Meyer presented a paper that specifies the data ownership concept as a foundation for the development of organizational structures and organizational rules for data warehousing, the data ownership concept is specified. Based on that concept, a two-dimensional organizational structure was presented that allows combining infrastructural competencies and content competencies and this concept was implemented in a large Swiss bank. They concluded as continuing research is needed to validate the concept by analyzing additional companies and industries and it has to be analyzed whether operational usage of information integration infrastructures leads to similar organizational concepts. [5]

Deis proposed a data warehousing systems that enabled enterprise managers to acquire and integrate information from heterogeneous sources and to query very large databases efficiently. Building a data warehouse requires adopting design and implementation techniques completely different from those underlying operational information systems. Though most scientific literature on the design of data warehouses concerns their logical and physical models, an accurate conceptual design is the necessary foundation for building a DW which is well-documented and fully satisfies requirements. In this paper they formalize a graphical conceptual model for data warehouses, called Dimensional Fact model, and propose a semi-automated methodology to build it form the pre-existing (conceptual or logical) schemes describing the enterprise relational database. In this paper they proposed a conceptual model for data warehouse design and a semi automated methodology for deriving it from the documentation describing the information system of the enterprise. Their work was devoted to developing the methodology for logical design and implementing it within an automated tool. [14]

The multidimensional model may be mapped on the logical level differently depending on the underlying DBMS directly supporting the multidimensional model was used, fact attributes were typically represented as the cells of multidimensional arrays whose indices were determined by key attributes[18] On the other hand, in relational DBMSs the multidimensional model of the DW is mapped in most cases through star schemes[15] consisting of a set of dimension tables and a central fact table. Dimension tables are strongly renormalized and used to select the facts of interest based on the user queries. The fact table stores fact attributes; its key is defined by importing the keys of the dimension tables. Different versions of these base schemes had been proposed in order to improve the overall performances [17], handle the data [20] and optimize the access to aggregated data[19] In particular, the efficiency issues raised by data warehousing have been dealt with the means of new indexing techniques among which we mention bitmap indices. [20]

A dimensional scheme consists of a set of fact schemes. The components of fact schemes are facts, measures, dimensions and hierarchies. A fact is a focus of interest for the decision-making process; typically, it models an event occurring in the enterprise world (e.g., sales and shipments). Measures are continuously valued (typically numerical) attributes which describe the fact from different points of view; for instance, each sale is measured by its revenue [14].

G.D.K. Kishorel, and et al described about the real challenge for tax agencies was to improve the tax collection process and to formulate optimized legislative polices to provide better taxpayer services. The existing Online Transaction Processing (OLTP) systems were not adequate to fulfill the need of complex tax collection analysis. And it is very difficult to come up with the optimized model. To make the best optimized model it has to be built around facts and it will be helpful to use a Tax Decision Support System to make strategic decisions. To solve this problem, this paper introduces the decision making activities and tax reduction methods for tax revenue. To maintain the effective tax management system, they introduced data warehouse and OLAP techniques. The goal of their paper was an ultimate tax data warehouse is to provide decision-makers access to consistent, reliable, and timely data for analytical, planning and tax assessment purposes that allows for easy retrieval, exploration and analysis. [2]

Robert Winter and et al presented a paper about the need to connect large numbers of decision support systems to large numbers of operational systems by providing a hub for subject-based, historical, consistent, and non-volatile information. By connecting decision support systems and operational systems to a (logically) centralized hub, the number of interfaces can be reduced dramatically and information quality can be guaranteed more effectively. [5]

Distributed data mining in data warehouse Distributed Decision Tree Algorithm is proposed in this paper. Decision trees are eligible for generating unambiguous and comprehensive rules from data stored in the data storage any by doing this to support the decision making – e.g. after constructing a decision tree from particular data storage the company managers are able to find optimal decisions for managing a company or its part much easier or they can predict the influence of their decisions [6]

Gilberto Montibeller and Alberto Franco delivered the paper which discussed the use of MCDA for supporting strategic decision making, particularly within strategy workshops and a framework to employ Multi-Criteria Decision Analysis for supporting strategy workshops. They explained the nature of strategic decisions and the characteristics of the strategic decision making process and tested the technical issues associated with the content of strategic decisions, and the social aspects that characterize the processes within which they are created. They gave suggestions on how to implement these proposals, and illustrate them with examples drawn from real-world interventions in which we have participated as strategic decision support analysts. They suggested that there were two main aspects that have to be addressed by the decision analyst, if they want to support strategic decision making processes. The first is related to content issues, in particular in dealing with epistemic uncertainty, multiple organizational objectives, complex policies and long-term consequences. The second aspect concerned process issues, in particular being an active listener, dealing with group dynamics and helping the group to reach closure [7].

Multi-criteria decision analysis (MCDA) is an umbrella approach that has been applied to a wide range of natural resource management situations. This paper aims to fulfill two purposes, the first one is to offer a critical review of MCDA methods applied to forest and other natural resource management and the second purpose is to describe new MCDA paradigms aimed at addressing the inherent complexity of managing forest ecosystems, particularly with respect to multiple criteria, multi-stakeholders, and lack of information. These new perspectives do not undermine the value of traditional methods; rather they point to a shift in emphasis from methods for problem solving to methods for problem structuring. They suggested that, this MCDA offers a suitable planning and decision-making framework for natural resources management. Because it is inherently robust, it can also provide a convenient platform that lends itself well in bridging the gap between the soft qualitative planning paradigm and the more structured and analytical quantitative paradigm. Approaches that integrate these two paradigms offer some promise in terms of more adequately accommodating the inherent complexity of natural resources management, embracing ecological, biophysical, and social components, and capturing the multitude of concerns, issues and objectives of stakeholders. [8]

3.Extraction, Transformation, and Loading (ETL):

Extraction, Transformation, and Loading (ETL) processes are accountable for the operations taking position in the back stage of data warehouse architecture. In a high level description of an ETL process, first the data are extracted from the source data stores, which can be in a relational and/or a semi-structured format. In typical cases, the source data stores can be On-Line Transactional Processing (OLTP) or legacy systems, Files under any format, web pages, various kinds of documents or even data coming in a streaming fashion. Typically, only the data that are different from the previous execution of an ETL process (newly inserted, updated, and deleted information) should be extracted from the sources. After this phase, the extracted data are propagated to a special–purpose area of the warehouse, called Data frequently used transformations include filters and checks to ensure that the data propagated to the warehouse respect business rules and integrity constraints, as well as schema transformations that ensure that data fit the target data warehouse schema. Finally, the data are loaded to the central data warehouse (DW) and all its counterparts (e.g., data marts and views). In a traditional data warehouse setting, the ETL process periodically refreshes the data warehouse during idle of low-load, periods of its operation (e.g., every night) and has a specific time-window to complete. Nowadays, business necessities and demands require near real-time data warehouse refreshment and significant attention is drawn to this kind of technological advancement [9].

ETL is a powerful metadata-based process that extracts data from sources systems and loads data into a data warehouse. In the process, it performs transformations designed to improve overall data quality and report ability. The metadata maintains a history of the transforms and provides this information to business users through data lineage and impact analysis diagrams [1].

Despite the fact that ETL took its name and separate existence during the first decade of the 21st century, ETL processes have been a companion to database technology for a lengthier period of time in fact, from the beginning of its existence. During that period, ETL software was just silently hidden as a routine programming task without any particular name or individual importance. ETL was born on the first day that a programmer constructed a program that takes records from a certain persistent file and populates or enriches another file with this information. Since than, any kind of data store that the original one is a form of an ETL program. Apart from this low profile programming task, research efforts have long hidden ETL tasks, although not much attention was paid to them. The earliest form of ETL system that we know of goes back to the EXPRESS system [13] that was intended to act as an engine that produces data transformations given some data definition and conversion nonprocedural statements. In later years, during the early days of data integration, the driving force behind data integration were wrapper-mediator schemes; the construction of the wrappers is a primitive form of ETL scripting [12]. In the mid '90's, data warehousing came in the central stage of database research and still, ETL was there, but hidden behind the lines. Popular books [11] do not mention the ETL triplet at all, although the different parts (transformation, cleansing, staging of intermediate data, and loading) are all covered.

U. Dayal and et al described that the software engineering community has measures for evaluating the quality of software design, so they adapted some to the quality of ETL designs. Due to space limitations, they focused on a subset of the aforementioned metrics, but for a description of other metrics they referred the interested reader to [17].

ETL workflows are much more complex than tradition related queries, thus the well-known techniques for multi-query optimization are not enough in this context. Still, they could leverage knowledge acquired from query processing. The rule that the most restrictive operations should be placed at the start of the flow, applies here as well. Such algebraic optimization can be done is several phases of the design, conceptual, logical, and physical [19,20,21]. The commercial ETL software either does not support any automatic optimization capabilities or offers limited optimization functionality (e.g., the Push Down optimization that pushes, usually small, portions of the ETL workflow, either its beginning or its end to the DBMS trying to leverage its optimization power [18]. Alkis Simitsis and et al, presented the QoX metric suite that aims at handling such metrics in all the ETL design levels and discussed the interrelationships and dependencies among the metrics that lead to tradeoffs for alternative optimizations of ETL process. Another challenge is creating tools to automate the optimization, which is a topic they were working on [16].

4.Proposed Work:

This paper integrates the concept of Hyper ETL and decision analysis methodologies to get a better decision in the business. The connotation of this Hyper ETL was examined through sample sales records and we initiated that and it took 89 minutes for nearly 15 lakhs records for which transformation time was less than the existing transformation time , in order to provide an optimal solution to the policy makers to take right decision at right time for producing turn over. The job obtainable in this work is planned at discover an effective decision making for sales promotion in sales data mart using hyper ETL.

In this paper, we incorporate the decision analysis methodology with output of Hyper ETL tool to make right decision in order to develop the sales promotion. Decision analysis is the discipline of evaluating complex alternatives in terms of values and uncertainty. Values are generally expressed monetarily because this is a major concern for management. Furthermore, decision analysis provides insight into how the defined alternatives differ from one another and then generates suggestions for new and improved alternatives. Numbers quantify subjective values and uncertainties, which enable us to understand the decision situation. These numerical results then must be translated back into wards in order to generate qualitative insight. A decision needs a decision maker who is responsible for making decisions. This decision maker has a number of alternatives and must choose one of them. The objective of the decision-maker has no control over may have occurred. Each combination of alternatives, followed by an event happening, leads to an outcome with some measurable value. Managers make decisions in complex situations. Decision tree and payoff matrices illustrate these situations and add structure to the decision problems [22].

Consider the following Sales Promotion Table

Places / Items	I1	I2	I3	I4	I5	I6
P1	30	20	10	40	50	60
P2	10	40	20	30	60	50
P3	20	30	30	10	40	50
P4	40	10	40	20	30	10
P5	30	20	50	10	40	60
P6	50	50	60	30	20	10

Figure: 1

Considering the uncertain environment, the chance that “good decisions” are made increases with the availability of “good information.” The chance that “good information” is available increase with the level of structuring the process of Knowledge Management. However, for private decisions one may rely on, e.g., the psychological motivations, as discusses under “Decision Making under Pure Uncertainty” in this site. Moreover, Knowledge Management and Decision Analysis are indeed interrelated since one influences the other, both in time, and space [23, 24, 25] .
Decision analysis consists of five criteria’s namely Minimum decision criteria ,Maximum criteria, Hurwitz criteria, La Placian criteria and MiniMax Regret decision analysis criteria and it is used to make better decision making . It provides the movement of sales to the particular places based on the decision analysis criteria

Criterion / Places	P1	P2	P3	P4	P5	P6
Minimum	10	10	10	10	10	10
Maximum	60	60	50	40	60	60
Hurwitz	35	35	30	25	35	35
Laplacian	52.5	52.5	45	37.5	52.5	55
MiniMax Regret	50	40	30	50	30	40

Figure: 2

Figure: 3

1)Maximum And 2)Minimum Criterion: From the above table

Maximum Quantity of sales	Minimum Quantity of Sales
60	10
60	10
50	10
40	10
60	10
60	10

Figure : 4

3) Hurwicz Criterion (Assign a = .5)

h = a X maximum + (1-a) X Minimum
P1(h) = .5 *60+(1-.5)*10
= 30.0 + .5 * 10
= 30+5
= 35

Sales based on Hurwicz Criterion
35
35
30
25
35
35

Figure: 5

4) Laplace Criterion (Equal value = ¼)

Places / Items	I1	I2	I3	I4	I5	I6	Laplace
P1	7.5	5	2.5	10	12.5	15	52.5
P2	2.5	10	5	7.5	15	12.5	52.5
P3	5	7.5	7.5	2.5	10	12.5	45
P4	10	2.5	10	5	7.5	2.5	37.5
P5	7.5	5	12.5	2.5	10	15	52.5
P6	12.5	12.5	15	7.5	5	2.5	55

Figure -6

P1(I1) = 30*1/4=7.5

5) Minimax Regret Criterion :

Places / Items	I1	I2	I3	I4	I5	I6
P1	30	20	10	40	50	60
P2	10	40	20	30	60	50
P3	20	30	30	10	40	50
P4	40	10	40	20	30	10
P5	30	20	50	10	40	60
P6	50	50	60	30	20	10

Figure 7

Places / Items	I1	I2	I3	I4	I5	I6	Maximum Regret
P1	20	30	50	0	10	0	**50**
P2	40	10	40	10	0	10	40
P3	30	20	30	30	20	10	30
P4	10	40	20	20	30	50	**50**
P5	20	30	10	30	20	0	30

Figure: 8

From the Minimax Regret Criteria We have to draw the Decision Tree. This is used to determine the Maximum and Minimum Quantity of Sales based on Places and Items.

From the Decision Analysis Criteria we take the following Four Decisions
* 1. Maximum and Hurwicz shows Same movement of Sales Quantity.
* 2. Minimum Criteria shows equal value of sales quantity. So we does not take this Criteria for Decision making
*3. Laplace shows One place only. It is also not possible to take decision.
*4. Minimax Regret Criteria shows two places. So It is better for taking decision on Sales Promotion

Items/ Criterion	Minimum	Maximum	Hurwitz	Laplacian	Regret
I1	10	50	30	45	40
I2	10	50	30	42.5	40

I3	10	60	35	52.5	50
I4	10	40	25	35	30
I5	20	60	40	60	40
I6	10	60	35	60	50

Figure : 9

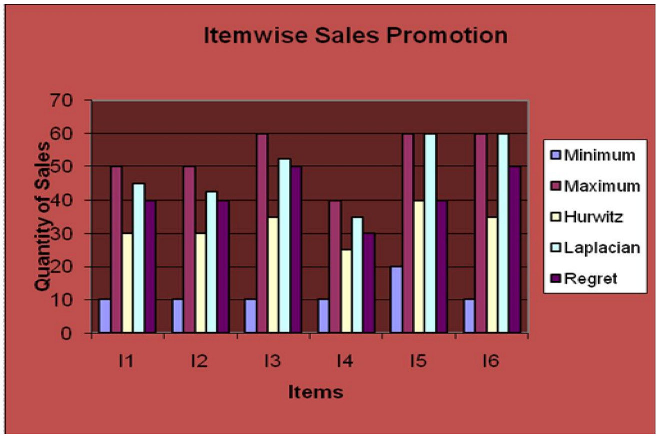


Figure: 10

* From the various Criteria we take the following Decisions.
1. Maximum Criteria - * P1,P2,P5,P6 shows the Best Movement of Sales.
2. Minimum Criteria - * P1,P2,P3,P4, P5,P6 all the Place shows the Best Movement of Sales.
3. Hurwicz Criteria - * P1,P2,P5,P6 shows the Best Movement of Sales.
4. Laplace Criteria - * P6 shows the Best Movement of Sales.

Good decisions need clear objectives. These should be specific, measurable, agreed, realistic and time-dependent. The following process might apply to the development of a policy, a project.

- Identifying objectives
- Identifying options for achieving the objectives
- Identifying the criteria to be used to compare the options
- Analysis of the options
- Making choices, and
- Feedback. [4]

An area of decision making closely related to the topics covered in this manual involves the use of decision trees to help identify good strategies for planning a response to a set of interdependent decisions sequenced through time. The actual outcome of each of the individual decisions at each stage is not known with certainly. Appropriate analysis of the tree allows the decision maker to develop, from the outset of the decision process, a contingent decision strategy. Decision trees have as their prime focus the question of uncertainty about the outcomes of decisions and. Rather; they reflect relatively simple appraisal guidelines, such as straightforward maximization of profit or minimization of cost [4].

5. Experimental analysis and Implementation:

We have presented an enchained approach of Hyper ETL to improve the concert sales data mart. To provide evidence for sales data mart which used Hyper ETL tool, trials were carryout in sales records by applying decision analysis criteria. And we found that we suggest the regret decision analysis produce, the apt choice for decision making. Experimental and Implementation results are given below.

Figure:11

About Items

Maximum	**Minimum**		**Maximum**	**Minimum**
P1 - 10	**50**	**I1**	**10,20**	**40**
P2 - 10,20	**50**	**I2**	**20**	**Nil**

P3 - 20,30,10	**50**	**I3**	**10,20,30**	**50**
P4 - 10	**40**	**I4**	**10,10**	**30**
P5 - 20,10	**50**	**I5**	**NIL**	**50**
P6 - 10	**30**	**I6**	**10,10**	**50,50**

Figure: 12

Maximum Quantity of sales
Based on Items

	I1	I2	I3	I4	I5	I6
P1	20	30	50	0	10	0
P2	40	10	40	10	0	10
P3	30	20	30	30	20	10
P4	10	40	20	20	30	50
P5	20	30	10	30	20	0
P6	0	0	0	10	40	50

Figure 13

Maximum Quantity of Sales based on Items
From this we take the decision
1. In I1 , P2,P3 move more Places than P4
2. In I2 , P5 move more Places
3. In I3 , P1,P2,P3 move more Places than P5
4. In I4 , P3,P5 move more Places than P6
5. In I5 , NIL (Does not have Minimum Quantity of Sales)
6. In I6 , P4,P6 move more Places than P2,P3

Minimum Quantity of Sales based on Items
From this we take the decision
1. In I1 , P4 Move Less Places than P2 and P3
2. In I2 , P5 (Only one place)
3. In I3 , P5 Move Less Places than P1,P2 , P3
4. In I4 , P6 Move Less Places than P3 and P5
5. In I5 , P1(Move only one Place)
6. In I6 , P2,P3 Move Less Places than P4 and P6

Figure 14

Minimum Quantity of sales Based on Places

	P1	P2	P3	P4	P5	P6
I1	20	40	30	10	20	0
I2	30	10	20	40	30	0
I3	50	40	30	20	10	0
I4	0	10	30	20	30	10
I5	10	0	20	30	20	40
I6	0	10	10	50	0	50

Figure 15

Figure 16

From the chart ,we get the following interpretation

Maximum Quantity of Sales based on Places

From this we take the decision

1. In P1 , I3 move more Places than I5
2. In P2 , I1 move more Places than I6
3. In P3 , I1,I3, and I4 move more Places than I5
4. In P4 , I6 move more Places than I1
5. In P5 , I2,I4 move more Places than I3
6. In P6 , I6 move more Places than I4

Minimum Quantity of Sales based on Places

From this we take the decision

1. In P1 , I5 move Less Places than I3
2. In P2 , I6 move Less Places than I1,I3
3. In P3 , I6 move Less Places than I1,I3,I4
4. In P4 , I1 move Less Places than I6
5. In P5 , I3 move Less Places than I2,I4
6. In P6 , I4 move Less Places than I6

Depending on the Regret Sales, Decision Tree can be drawn

Figure 17

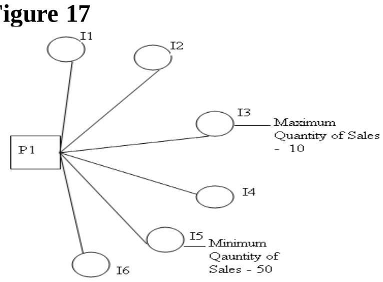


Figure 18

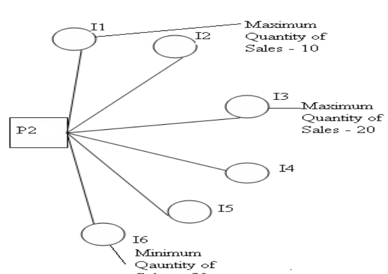


Figure 19

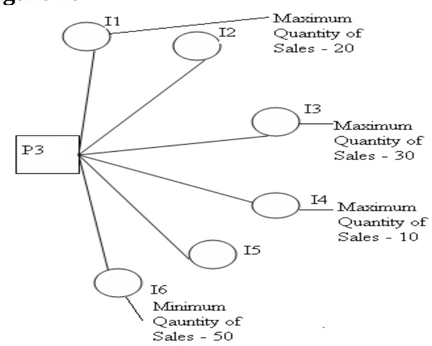


Figure 20

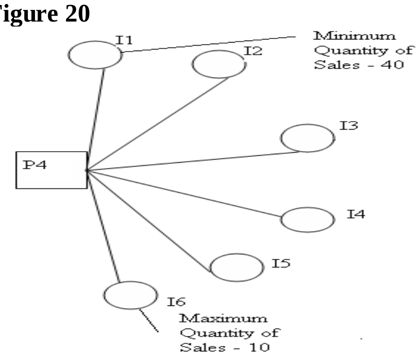


Figure 21

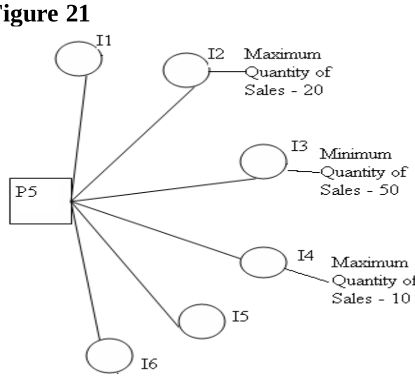
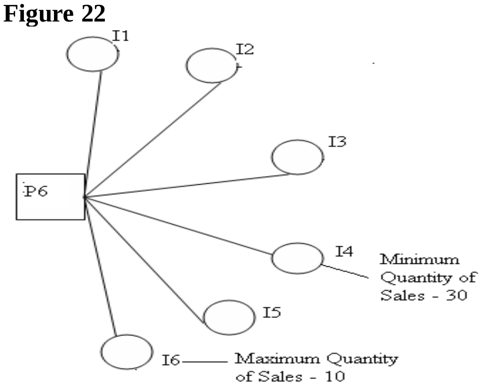


Figure 22



From the Decision Tree we can conclude that the minimum and maximum quantities of items sold in all places. Depending on the Decision Tree Minimum Quantity of product sold at maximum level in all places. With the help of this table, the sales manager can take decision to increase the sales promotion in a particular place.

	Maximum (sales)	Minimum(sales)
P1	10	50
P2	10,20	50
P3	20, 30, 10	50
P4	10	40
P5	20,10	50
P6	10	30

6. Conclusion:

We have presented the refined design of Hyper ETL which accomplishes enhances show of ETL, through reducing the data transformation time and cost and improves the throughput. In this paper, we made a amalgamation of the contribution of enhanced Hyper ETL Tool with decision analysis methodologies This decision analysis methodologies suggested the optimal solution and this holds great potential for dramatic business benefits and also provides decision makers access to consistent, reliable and timely data still much more to explore and enhance and efficient decision making to hike the sales promotion. The effort presented in this paper is intended to get better presentation of ETL progression deal with an effective decision making.

References

[1.] Master Data Management

[2.] Jeffrey R. Bocarsly, Ph.D Director – Functional Test Automation Group Complex ETL Testing; A Strategic Approach

[3.] DATA WAREHOUSING AND DATA MINING – A CASE STUDY Milija SUKNOVIC, Milutin CUPIC, Milan MARTIC – Trizon Group, Belgrade, Serbia and Montenegro Yugoslav Journal of Operations Research 15 (2005), Number 1, 125 – 145

[4.] Multi-criteria analysis: a manual

[5.] Organization of data warehousing in large service companies – A matrix approach based on data ownership and competence centers Robert Winter and Markus Meyer

[6.] Distributed Data Mining and Data Warehouse – Szappanos Tibor[1], Zolotova lveta[2], Landryova, Lenka[31]

[7.] Multi-criteria Decision Analysis for Strategic Decision Making Gilberto Montibeller and Alberto Franco

[8.] Multi-criteria decision analysis in natural resource management:- A critical review of methods and new modeling paradigms – G.A. Mendoza,*, H. Martin b

[9.] EXTRACTION, TRANSFORMATION, AND LOADING Panos Vassiliadis, University of loannina, loannina, Hellas, Alkis Simitsis, IBM Almaden Research Center

[10.] Jiwei, H., and Micheline, K., Data Mining: Concepts and Techniques, Simon Fraser University, 2001.

[11.] W.Inmon. Building the Data Warehouse. John Wiley & Sons, 2nd edition, 1996.

[12.] M.T. Roth and P.M.Schwarz. Don't scrap it, wrap it! A wrapper architecture for legacy data sources. In VLDB, pages 266 {275, 1997.

[13.] N.C.Shu, B.C. Housel, R.W. Taylor, S.P.Ghosh, and V.Y.Lum. Express: A data extraction, processing, and restructuring system. ACM Trans. Database Syst., 2(2): 134} 174,1977.

[14.] Decision making and Reduction Techniques for Tax Revenue using data Warehouse G.D.K.Kishore 1, B.Anusha 2, G.Poornima 3, G.Anusha4, K.Tejaswi5

[15.] Organization of data warehousing in large service companies – A matrix approach based on data ownership and competence centers Robert winter and Markus Meyer

[16.] QoX-Driven ETL Design: Reducing the Cost of ETL Consulting Engagements Alkis Simitsis, HP Labs, Palo Alto, Ca, USA, ets.

[17.] U. Dayal, M. Castellanos, A. Simitsis, K. Wilkinson. Data Integration Flows for Business Intelligence. In EDBT, 2009.

[18.] Informatica. How to Achieve Flexible, Cost-effective Scalability and Performance through Pushdown Processing. White Paper, November 2007.

[19.] N.N.Dalvi, S.K.Sanghai, P.Roy, S.Sudarshan. Pipelining in Multi-Query Optimization. In PODS, 2001.

[20.] P.Roy, S.Seshadri, S.Sudarshan, S.Bhobe. Efficient and Extensible Algorithms for Multi Query Optimization. In SIGMOD, 249-260, 2000.

[21.] T.K.Sellis. Multiple-Query Optimization. In ACM Trans. Database Syst. 13 (1), pp.23-52,1988.

[22.] Manning N., et al., Strategic Decision Making in Cabinet Government: Institutional Underpinnings and Obstacles, World Bank, 1999.

[23.] Berger J., Statistical Decision Theory and Bayesian Analysis, Springer, 1978. Corfield D., and J. Williamson, Foundations of Bayesianism, Kulwer Academic Publishers, 2001. Contains Logic, Mathematics, Decision Theory and criticisms of Bayesianism.

[24.] Press S., and J. Tanur, the Subjectivity of Scientists and the Bayesian Approach, wiley, 2001. comparing and contrasting the reality of subjectivity in the work of history's great scientists and the modern Bayesian approach to statistical analysis.

[25.] Tanaka H., and P.Guo, Possibilistic Data Analysis for Operations Research, Physica – Verlag, 1999.