Data Warehouse Development Standardization Framework (DWDSF): A Way to Handle Data Warehouse Failure

Deepak Asrani[1], Dr. Renu Jain[2], Dr. Usha Saxena[3]

1(Department of Computer Science Engineering, Teerthanker Mahaveer University, Moradabad, U.P (India))
2(Department of Computer Science Engineering, University Institute of Engg. & Tech, Kanpur, U.P (India))
3(State Institute of Health and Family Welfare, Indira Nagar, Lucknow, U.P (India))

Abstract: *Why does large number of data warehousing projects fail? How to avoid such failures? How to meet out user's expectations and fulfil data analysis needs of business managers from data warehousing solutions? How to make data warehousing projects successful? These are some of the key questions before data warehouse research community in the present time. Literature shows that large numbers of data warehousing projects undertaken eventually result in a failure. In this paper, we have designed a framework named "Data Warehouse Development Standardization Framework" (DWDSF), to help data warehouse developer's community in implementing effective data warehousing solutions. We have critically analysed literature to find out possible reasons of data warehouse project failure. Our framework has been designed to overcome such issues and enable implementation of successful data warehousing solutions. To verify usefulness of our framework, we have applied guidelines of DWDSF framework to design and implement data warehousing solution for National Rural Health Mission (NRHM) project which offers various health services throughout the country. The developed solution is giving results for all type of queries business managers want to run. We have shown results of some sample queries executed over the implemented data warehouse repository. All results are meeting out business manager's query expectations.*

Keywords: *Data warehouse failure, Data warehouse framework, Data warehousing practices, Operational System limitations*

I. Introduction

Any business organization having quality operational system implemented is usually able to generate useful MIS reports to visualize and manage its business processes and operations smoothly. But with the help of operational system, business managers of the organization cannot generate reports useful for long term analysis of their business processes and strategic decision making. Operational system reports do not have features to produce deep insights of business operations to support strategic decision making. To meet out such requirements one needs to have a quality data warehousing solution. From literature it is seen that major focus of data warehousing research has been on improving physical implementation and run time query execution performance. Limited attention is paid on issues related to data warehousing practices. Data warehousing solution developers can ensure effectiveness in the developed solutions only if they follow a systematic approach of data warehouse development procedure based on some well defined and tested framework that not only offers specific guidelines about steps to be followed for data warehouse design and implementation but also suggests the approach to be used during each step. We have designed a framework named "Data Warehouse Development Standardization Framework" (DWDSF), to help data warehouse developer's community in implementing effective data warehousing solutions. We have experimentally verified its usefulness. For dealing with issues of data warehouse failure, we first need to define and understand clearly what actually is meaning of data warehouse failure. By data warehouse failure we mean not meeting out specified goals and objectives of the organization from the developed solution. Research literature on data warehouse failure identifies types of failures broadly categorized in two types. i. Failures due to managerial aspects and ii.Failures due to technical aspects. Table 1 lists out some of the managerial aspects of data warehouse failure and Table 2 lists out some of the technical aspects of data warehouse failure.

Table 1: List of managerial aspects of data warehouse failure

Sr.No.	Managerial Aspects of Data warehouse Failure
1	Running out of estimated budget
2	Running beyond planned schedule of solution implementation
3	Solution not able to meet out user expectations
4	Project turns out to be unduly costly affair
5	Developed solution is not able to offer expected benefits to the organization

Table 2: List of technical aspects of data warehouse failure

Sr.No.	Technical Aspects of Data warehouse Failure
1	Expected functionality not implemented
2	Unsatisfactory performance
3	Availability issues
4	Poor scalability of solution
5	Reports not fully dependable
6	Usability issues

If we closely analyze reasons of data warehousing solution's failure, we find that most of the reasons cited are related to issues of project development process (procedural issues). It is because efficient tools and technologies are available at our disposal. We only need to rightly identify and follow a systematic approach of data warehouse development procedure to ensure successful implementation of developed solution. We also need to ensure utilization of available tools and technologies correctly. Planned development procedure is one which is based on proper documentation, is well communicated, considers specific list of user expectations, and is well estimated in terms of time and expenditure.Application development refers to computerization of functioning of an organization or a system. Functioning is routine business activity which is almost always clear and can be easily defined by the users. It is therefore easy to ensure its successful implementation. On the other hand data warehousing project implementation refers to developing a platform of data repository to serve business managers with variety of complex business queries for in depth analysis of all business subjects. In the case of data warehouse implementation, requirements are generally not well defined by the users and therefore needs to be determined carefully to succeed in implementation. Application development and data warehouse development both need software engineering approach, project management skills, and tools and techniques but with a different approach. Programming projects are easy to implement. Why? Because there we need to focus only on what is to be done. How it is to be done, is taken care by inbuilt code, which is pre tested for quality. In data warehousing projects, it becomes important to determine how things are to be done due to unavailability of standardized approach. We have tried to standardize the procedure of data warehousing to improve success rate of data warehousing projects.

Nature of queries of operational system and data warehousing are different from each other. Operational system queries report business data which is current or of recent past, and is generally atomic in nature. On the other hand, data warehousing queries report result of business process analysis over a long period of time span either in a single time slot or in a sequence of multiple time spans. In the absence of data warehousing solution, performance assessment of various services provided, and their effects are collected by means of carrying out large scale public surveys to get public feedback. In case of data warehousing solution, getting such information becomes automatic with the help of running suitable database queries over developed data warehousing solution. These are major benefits of implementing data warehousing solutions.

Organization of this paper is as follows. Section 2 covers literature review. Section 3 discusses design of data warehouse development standardization framework (DWDSF). Section 4 gives details of National Rural Health Mission (NRHM) project and its data warehousing solution. Section 5 includes some of the business manager's query expectations, SQL queries for their requirements and their results. Section 6 concludes the paper.

II. Literature Review

Data warehouse failure can be attributed to issues of data quality, database schema, quality of modeling tool, and quality of operational system [4]. Misunderstood user expectations, lacking in understanding the data, adopting an over complicated architecture for the solution, considering user defined requirements as final and complete, and always blaming quality of data for data warehouse failure are key reasons of data warehouse project's failure [6]. [3] Highlighted the problem of lack of standardization of data warehouse development methodologies. Organizations often struggle with successfully delivering a data warehouse solution due to execution challenges, and poor planning which could lead to years of extended wait times before organization can realize the benefit of the data warehouse [6]. Gartner Says More Than 50 Percent of Data Warehouse Projects Will Have Limited Acceptance or Will Be Failures Through 2007 as a result of a lack of attention to data quality issues, according to Gartner, Inc [2]. "Consistency and accuracy of data is critical to success with BI, and data quality must be viewed as a business issue and responsibility, not just an IT problem" [2]. "It is hard to believe that IT organizations still build data warehouses with little or no business involvement," said Frank Buytendijk, research vice president at Gartner"[2]. Failure rate of data warehousing projects has been a major concern in data warehousing industry. Data warehousing is an essential element of decision support. It aims at enabling the knowledge user to make better and faster daily business decisions [1]. The process of data warehousing starts by analyzing operational system of the organization. Incomplete and poor operational system design is bound to put hurdles in smooth design and implementation of data warehousing solution [7].

Research contributions in the field of data warehousing could be classified in the following two broad categories.

1. Developing tools and technologies to support data warehouse implementation.
2. Developing methodologies and procedural steps to be used by data warehouse developers to help them in effective implementation of data warehousing solutions.

III. Design of Data Warehouse Development Standardization Framework (DWDSF)

For accomplishing any challenging task following factors are important: i. Availability of advance tools and technologies. Ii. Knowledge of appropriate approach to be followed. If we consider implementation of data warehousing solutions, we find that advance tools and technologies are available. We only need to identify and follow the approach that would be effective for successful implementation of data warehousing solution. We have designed a framework named "Data Warehouse Development Standardization Framework" (DWDSF), to help data warehouse developer's community in implementing effective data warehousing solutions. Our framework (DWDSF) proves highly useful in helping out in deciding on such issues of appropriate approach selection. DWDSF framework specifies i. Steps of data warehouse development process, ii.Decisions to be taken during each step, and iii.Approach to be followed, during each step.

Data warehouse development steps proposed in our framework take care of managerial aspects of data warehouse implementation, whereas details of activities proposed during each of the steps in the framework take care of technical issues of data warehouse implementation.

Our framework's methodology is based on two layer approach. Layer 1 describes macro steps of the data warehouse development process. Layer-2 describes micro steps under each macro step and its determinants. Layer-1 develops general understanding of data warehouse development approach to be followed whereas layer 2 ensures that each of the steps is undertaken with a clear idea of its purpose. Two-layer approach ensures that we strictly adhere to the proposed framework and at each of the micro step check if we are getting the mentioned outcome of that micro step. Macro layer guides about broad steps to be followed for data warehouse development where as micro layer of the framework guides about specific atomic level activities to be performed during each of the macro step and describes its purpose in terms of outcome. Our framework is "what and how" kind of framework which not only tells the steps to be followed but also describes the way to do each step. The major focus of our framework is on requirement analysis and data warehouse design phase, because these two are highly complex tasks in data warehousing projects. Implementation of designed system is relatively simple, involving routine activities of physical implementation. Implementation involves execution of ETL code or steps to clean and load data into data warehouse schema.

Our framework suggests following basic steps (Macro Steps) to be followed for data warehouse implementation:

1. Understand goals and objectives
2. Identify measurements to be made that could assess if goals and objectives are met
3. Investigate data items required for above measurements
4. Design schema to be populated with the identified data items
5. Iterate above steps till complete system is designed.
6. Implement designed system and populate with data.

Many research papers have repeatedly given steps to be followed for data warehouse solution development but internal details of steps are very brief. We have tried to dig deep into Micro details of steps that are useful for data warehouse developer's community. We have given details of the micro layer of the framework in Table 3 which are represented in the form of a Purpose table for data warehouse development process explaining purpose of each atomic step.

Table 3: Purpose table for data warehouse development process explaining purpose of each atomic step

Sr. No.	Data warehouse development atomic step	Purpose of the atomic step
1.	Identify Goals and Objectives	Determines reporting needs
2.	Identify Key Business Processes	Determines facts to be measured and dimensions of analysis
3.	Identify Organizational Structure	Determine different business users
4.	Investigate functionality supported in the operational system	Determines quality of the operational system Two-way quality assessment is performed i. At its own, ii. Quality assessment for data warehouse implementation
5.	Analyze reports generated by the operational system	Determines quality of the operational system
6.	Identify reports the operational system could not generate	Determines limitations of the operational system
7.	Interact with different users groups	Determines quality of operational system, users expectations from data warehousing solution

8.	Analyze operational system application forms and database schema	Determines data items available in the operational system and data items that are missing
9.	Analyze query types business managers want to run	Determines data warehouse schema design details, and grain of measurements to be taken in data warehouse schema design
10.	Analyze parameters of business process analysis	Determines dimension tables and their attributes
11.	Analyze key time intervals to be taken in most of the reporting	Determines execution schedule of ETL for population and updating of data warehouse schema
12.	Design dimensional model and design ETL code	Determines dimension tables, fact tables, their attributes, and schedules and inputs of ETL code execution
13.	Create tables, run ETL code,	Completes data warehouse schema implementation
14.	Run queries to generate reports to test data warehouse	Tests data warehouse for its performance and determines its quality

Data warehouse development challenges are due to multiple data sources, multiple data technologies, high volume, lack of user awareness about the data warehouse technology and its uses, design decisions for facts and dimension tables, about de-normalization or snow-flaking or fact table constellation, decision of data grain selection, architecture selection for implementing data warehouse, selection of suitable materialized views, design of ETL code, and complex data warehouse requirement analysis.

"Problem is not with the technology. It is with the approach we follow". Therefore our framework is based on procedural aspect of data warehousing. It is related to the subject of software engineering therefore we have taken standard SDLC and have tried to customize it for data warehousing projects.

Techniques applied for system investigation:

In our framework we have proposed following methods of system investigation:

1. Questionnaire distribution to business managers to collect their response.
2. Personal interview of business managers.
3. Organizational document study.
4. Analysis of operational system input forms and output reports.
5. Users query expectations.

Basic steps of DWDSF are given in table no. 4.

Table 4: Basic steps of DWDSF

Sr. No.	Data warehouse Implementation Steps
1.	Preliminary Investigation
2.	Assessment of Operational System
3.	Feasibility test report
4.	Analysis of needs
5.	Data warehouse design
6.	Data warehouse implementation

3.1 Features of DWDSF Framework

1. Integrated approach has been developed which guides in every aspect of data warehousing practices or data warehouse project execution.
2. The proposed Framework is based on procedural aspects of data warehousing that helps data warehouse developers to follow specific approach and ensure expected outcome from any project
3. Knowledge of physical implementation and the physical implementation methods used by a specific data warehousing tool helps in deciding which tool to use in a particular case but cannot govern developers about what steps to follow and what approach to follow during each step. This is where our framework proves highly useful.

3.2 Fitness test of operational system and its significance in data warehouse implementation

Signs of failure are "deciding to proceed for data warehousing with unfit operational system". Whereas, very first sign of success is "deciding to proceed for data warehousing only if operational system is fit for data warehousing". Taking decisions of data warehouse implementation without provisioning for current limitations leads to an attempt that most likely fails in the end. If operational system is not fit for data warehousing, we can suggest improvements in the operational system and postpone the data warehouse implementation. If only to fulfil commercial interests of development organisation, the project is undertaken, it is likely to fail. We have worked to investigate fitness of operational system for data warehouse implementation, and offer details of improvements required in the operational system if it is not fit for the data warehousing.

3.3 Causes of Heterogeneity
1. Non uniform data formats or structures
2. Change of technology with time

Non uniform data formats or structures are manageable with precautions in operational systems whereas changes in technology with time are unavoidable and need to be managed during data warehouse design. No prior research suggests in the direction of improving operational system design improvements. Most of the works are based on basic assumption that operational systems in use are fit for data warehouse implementation. Our approach of work is different from others. Our Framework suggests that the first step of data warehouse implementation is assessment of operational system.

3.4 Feasibility analysis for Data warehousing
1. Is operational system strong enough to support data warehousing? yes or no if no, then feasibility test result is big no, for data warehousing.
2. Need of analytical assessment of business? yes or no if answer is no then feasibility test is big no.

Managing multiple technologies of data sources is a requirement for data warehouse implementation. One method to deal with this requirement is to use data format standardization standard called XML. Diagrammatic representation of the model for heterogeneity management is as follows:
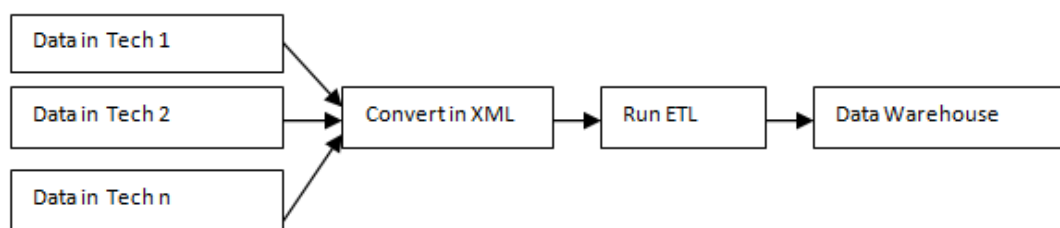

Fig No.-1: dealing with data heterogeneity using XML data standard.

IV. National Rural Health Mission and its Data warehousing solution

NRHM started with HMIS reporting system and later introduced MCTS application for mother and child care programme but without considering National level workload and infrastructure requirements. IT implementation goals and vision was not clearly defined.

NRHM goals for mother and child care module are as follows:
1. Reduce mother mortality rate (MMR)
2. Reduce infant mortality rate (IMR)
3. Reduce total fertility rate (TFR)
4. Prevention and reduction of anaemia in women aged between 15 to 49 years
5. Increase institutional delivery

MMR, IMR, TFR, prevention of anaemia, institutional delivery all are related to pregnant women health care during and post pregnancy. From the partial goal list, we are able to identify that we need to maintain data related to pregnant women registration and routine checkups details during and post pregnancy. By referring the health domain terminology from available organisational documents, we have identified following business processes
1. pregnant women ante natal care (ANC)
2. pregnant women delivery
3. pregnant women post natal care (PNC)

From these business processes, tentative facts are mother registration, mother checkups, mother delivery, mother PNC checkups details.

In addition to HMIS, and MCTS, NRHM project implemented State level websites which are maintained by each state separately.

We thoroughly investigated all operational systems implemented for NRHM and identified some useful points of concern. In depth investigation revealed that HMIS is a reporting system from facility centres to controlling body in a specific file format. It is actually a mix of manual data collection, compilation and computerized reporting platform and is not a full-fledged operational system. It involves both manual and computerized intervention to produce aggregated state level and national level reports. HMIS was developed to monitor various health programs for which fund is being released to all states, districts and so on but it is not able to support data warehouse designing, primarily because it does not maintain atomic data. It does not contain all data items which are required to analyze its business operations for strategic decision making. HMIS does not support any business process implementation and also is not able to generate any report on user request. It

does not have any user interface to take requests and generate customized reports. It is just a data collection platform from Central monitoring and evaluation division at national level, which then compiles the collected data and publishes compiled reports that are for user reference at any level.There are many data items which are not important for operational system but are very useful and are required for performance analysis. Such data items need to be included in operational system to facilitate effective data warehousing solution development. For example age of pregnant women or patient, address, socio economic status and other demographical data elements are not useful for operational system but are highly useful for demographic evaluation of operation's performance. Operational system needs or requirements and analytical system's needs or requirements are quite different. But analytical systems are developed from operational systems therefore needs for analytical system also need to be investigated for operational system development, to facilitate analytical application development of future. It is generally not considered, resulting in difficulties faced during development of analytical solutions like data warehouses. Data is not available for cases that received services from outside agencies (absence of external data).

Although, the operational system design is one of the most critical factor deciding success or failure of a data warehousing solution but most of the data warehousing research directly considers designing data warehouse without focusing much on operational system design and implementation. We have considered a step of quality assessment of operational system and its role in data warehouse designing. We have also discovered its solution which is based on a unique idea of designing operational system and Data warehousing Solution in an integrated manner to take care of objectives of both types of systems. We have followed an approach to first identify how much satisfied users are with the existing operational system and what problems they have been facing with their current operational system and why should they go for Data warehousing Solution and whether they need a data warehousing solution or need to improve their operational system first.Organizations size is big, volume of transaction is big, the age of project is also sufficient, need of complex analysis is also required. All above findings pass feasibility test of data warehousing implementation. But if the operational system is not systematically implemented it needs to be analyzed and improved so that, future Data warehousing Solution will also be a successful implementation. Why to design a Data warehousing Solution that will fail in future? First step should be to identify modifications required in the operational system for data warehouse design and re-design and implement it accordingly instead of going for a Data warehousing Solution. Can we say a project like NRHM does not require Data warehousing Solution? if no, is NRHM project's computerization in the current form suitable to support data warehouse implementation? if no., then what is next solution or option? Answer is to investigate operational system of NRHM, then do the required implementation changes in the operational system and then either sequentially or in parallel, proceed for data warehousing implementation. If we you look into Indian governments scenario, most of the projects despite being at National level are not fully automated. In such situations, real reporting is not possible. It in turn would result in decision taking which is not based on real facts. Why government schemes prove inefficient? Because, they are not able to pin point the problems in the absence of effective operational system.

Table 5: Information useful in size estimation

Sr. No.	Element Description	Element's Count at National Level
1.	Community Health Centres (CHCs)	5396
2.	Primary Health Centres (PHCs)	25308
3.	Health Sub Centres (SCs)	153655
4.	States	36
5.	Doctors	23089
6.	Ayush Doctors	11925
7.	Ayush Paramedicos	4785
7.	Auxillary Nurse Midwife (ANM)	70891
8.	Accredited social health activists (ASHA)	890000

Table 5 contains information useful in size estimation of data warehousing solution for NRHM project.
HMIS focused on reporting figures without considering parameters responsible for these figures. MCTS also has some of the required parameters missing. We have implemented ETL simulation code for generating various measurement counts based on inputs received from business managers and populating data warehouse schema. It was done due to non availability of atomic data in the operational system. Sample simulation code is given in figure number 2.

```
import java.sql.*;
public class ANCRegistrationPopulate {
    static final String JDBC_DRIVER = "oracle.jdbc.driver.OracleDriver";
    static final String DB_URL = "jdbc:oracle:thin:@localhost:1521";
    static final String USER = "d1";
    static final String PASS = "d1";
```

```java
public static void main(String[] args) {
    int darc_Id=1;
    int loopStart=1;
    int loopEnd=3941;
    Connection conn = null;
    PreparedStatement prSt = null;
    int k = 0;      //m_h_p
    int l = 0;      //m_p_p
    int m = 0;      //m_s_e_s
    int n = 0;      //anm
    int p = 0;      //asha
    int q = 0;      //registration_count
    int recordCount=0;
    try {
        Class.forName("oracle.jdbc.driver.OracleDriver");
        System.out.println("Connecting to a selected database...");
        conn = DriverManager.getConnection(DB_URL, USER, PASS);
        System.out.println("Connected database successfully...");
        System.out.println("Inserting records into the table...");
String sql = "INSERT INTO ANC_Registration VALUES(?,?,?,?,?,?,?,?,?)";
        prSt = conn.prepareStatement(sql);
        for (int i = loopStart; i <= loopEnd; i++) {
        for (int j = 1; j <= 9687; j++) {
            prSt.setInt(1, darc_Id);
            darc_Id++; //check shift appropriately
            prSt.setInt(2, i);
            prSt.setInt(3, j);
            while (!(k >= 1 && k <= 3620)) {
                k = (int) (Math.random() * 36200);
            }
            prSt.setInt(4, k);
            k = 0;
            while (!(l >= 1 && l <= 1222)) {
                l = (int) (Math.random() * 12220);
            }
            prSt.setInt(5, l);
            l = 0;
            while (!(m >= 1 && m <= 292608)) {
                m = (int) (Math.random() * 2926080);
            }
            prSt.setInt(6, m);
            m = 0;
            while (!(n >= 1 && n <= 113696)) {
                n = (int) (Math.random() * 1136960);
            }
            prSt.setInt(7, n);
            n = 0;
            while (!(p >= 1 && p <= 113696)) {
                p = (int) (Math.random() * 1136960);
            }
            prSt.setInt(8, p);
            p = 0;
            if ((j > 0 && j <= 1489) || (j > 1495 && j <= 1516) || (j > 2892 && j <= 2895) || (j > 2897 && j <=
8797)) {
                q = (int) (Math.random() * 40);
            } else if ((j > 1489 && j <= 1495) || (j > 2534 && j <= 2544)) {
                q = (int) (Math.random() * 120);
            } else if ((j > 2639 && j <= 2875) || (j > 9235 && j <= 9668) || (j > 1578 && j <= 2533)) {
                q = (int) (Math.random() * 300);
```

```
            } else if ((j > 1516 && j <= 1573) || (j == 2534) || (j > 2544 && j <= 2546) || (j > 2895 && j <=
2897) || (j == 9235) || (j == 9687) || (j > 8797 && j <= 8811)) {
                q = (int) (Math.random() * 900);
            } else if ((j > 1573 && j <= 1578) || (j > 2546 && j <= 2575) || (j > 2575 && j <= 2639) || (j > 2875
&& j <= 2872) || (j > 9668 && j <= 9686)) {
                q = (int) (Math.random() * 2500);
            } else if ((j > 8811 && j <= 9234)) {
              q = 0;
            } else {
              q = 1;
            }
            prSt.setInt(9, q);
            q = 0;
            prSt.executeUpdate();
            recordCount++;
          }
        }
      System.out.println("Inserted "+recordCount+"records into the table...");
    } catch (SQLException se) {
      se.printStackTrace();
    } catch (Exception e) {
      e.printStackTrace();
    } finally {
      try {
        if (prSt != null) {
          conn.close();
        }
      } catch (SQLException se) {
      }// do nothing
      try {
        if (conn != null) {
          conn.close();
        }
      } catch (SQLException se) {
        se.printStackTrace();
      }//end finally try
    }//end try
    System.out.println("Goodbye!");
  }//end main
}
```

Figure no. - 2, simulation code for generating various measurements counts

V. Results

For testing implemented solution, we tried to run queries requested by the business managers. We executed many queries requested by business managers and could generate the required reports as per user expectations. Some sample requested reports, executed SQL query, and the corresponding output is given as follows:

Query Objective:

Display number of women registered for ANC based on mother's education and family income.

SQL Statement:

Selectmother_education,family_income,sum(total_number_of_women_register)
"number_of_women_registered" from anc_registration b, mother_socioeconomic_status a where b.MOTHER_SOCIO_ECONOMIC_STATUS_I=a.MOTHER_SOCIOECONOMIC_STATUS_ID group by mother_education,family_income;

Output Result:

MOTHER_EDUCATION	FAMILY_INCOME	number_of_women_registered
Illiterate	Less Than Five Hundred	26496005
Illiterate	Five Hundred to Two Thousand	26252838
Illiterate	Two Thousand to Five Thousand	26231282
Illiterate	Five Thousand to Ten Thousand	26193811
Illiterate	More than Ten Thousand	26046384
Literate without School	Less Than Five Hundred	65100652
Literate without School	Five Hundred to Two Thousand	64831334
Literate without School	Two Thousand to Five Thousand	64777949
Literate without School	Five Thousand to Ten Thousand	64647611
Literate without School	More than Ten Thousand	65008987
Primary	Less Than Five Hundred	65110166
Primary	Five Hundred to Two Thousand	64611857
Primary	Two Thousand to Five Thousand	64610229
Primary	Five Thousand to Ten Thousand	64900781
Primary	More than Ten Thousand	64749036
Junior High School	Less Than Five Hundred	38386477
Junior High School	Five Hundred to Two Thousand	38623918
Junior High School	Two Thousand to Five Thousand	38391916
Junior High School	Five Thousand to Ten Thousand	38568355
Junior High School	More than Ten Thousand	38783664

Query Objective:

Display number of women registered for ANC based on mother's age range.

SQL Statement:

select age_range,sum(total_number_of_women_register) "TOTAL_NUMBER_OF_WOMEN_REGISTERED"from anc_registration b, mother_health_profile a where b.mother_health_profile_id=a.mother_health_profile_id group by age_range;

Output Result:

AGE_RANGE	TOTAL_NUMBER_OF_WOMEN_REGISTERED
Fifteen to Ninteen	358787552
Twenty to Twenty Four	356999924
Twenty Five to Twenty Nine	275261056
Thirty to Thirty Four	248242433
Thirty Five to Thirty Nine	126079298
Forty to Forty Four	59300773

Query Objective:

Display number of women registered for ANC based on distance from home to facility centre.

SQL Statement:

select distance_from_home_to_facility,sum(total_number_of_women_register)" TOTAL_NUMBER_OF_WOMEN_REGISTER" from anc_registration b, mother_health_profile a where b.mother_health_profile_id=a.mother_health_profile_id group by distance_from_home_to_facility; Output Result:

DISTANCE_FROM_HOME_TO_FACILITY	TOTAL_NUMBER_OF_WOMEN_REGISTERED
With in Three Kilometres	284720925
Between Three to Six Kilometres	284912481
Between Six to Ten Kilometres	285097249
Between Ten to Twenty Kilometres	285036463
More than Twenty Kilometres	284903918

Query Objective:

Display number of women registered for ANC based on year wise registration on each facility centre.

SQL Statement:

select a.facility_centre_name,c.year_of_Day,sum(total_number_of_women_register) from anc_registration b, facility_centre_dimension a, Date_Dimension c where b.facility_centreid=a.facility_centre_id and c.date_id=b.date_id group by facility_centre_name,c.year_of_Day;

Output Result:

FACILITY_CENTRE_NAME	YEAR_OF_DAY	TOTAL_NUMBER_OF_WOMEN_REGISTERED
Karaiyur	2005	40028
Athimanjeripettai	2005	39944
Negamam	2005	42282
Nallattipalayam	2005	42959
Madukkarai	2005	41086
Vadalur	2005	42920
Thoppur	2005	40247
Theerthamalai	2005	40320
Acharapakkam	2005	40047
Velliyanai	2005	40983
Manmangalam	2005	41311
NPHC Goberaha	2006	7435
Salt	2006	7023
Bahadrabad	2006	7255
Yameshwor	2006	6963
Rikhnikhal	2006	7278
Dugdda	2006	7328
Badalu	2006	7304
Chamba	2006	7179
Dwida	2006	7251
Almora Sub-divisional Hospital	2006	638119
Irumpedu	2007	7325
Desur	2007	7002
Chithathoor	2007	7033
Veppalodai	2007	6734

We have designed and implemented data warehouse on the basis of analysis of only one program of NRHM project out of 17 programs it is running. It is like a prototype which verifies quality of the developed the solution, gets feedback of user group to move further with design and development of data warehouse for other programs as well. The results generated by our solution were highly satisfactory. We could meet out all of user's reporting expectations which is highly encouraging result.

VI. Conclusion

Review of literature on data warehousing revealed issues of high failure rates of data warehousing projects. In this paper, we analyzed possible reasons of data warehousing project failure and found that most of the reasons are related to procedural aspects. We designed a framework called Data Warehouse Development Standardization Framework (DWDSF) which is based on procedural aspects of data warehousing, has tried to standardize steps of data warehousing and elaborates ways to undertake each of the steps. We have also discussed role of operational system in data warehousing and its feasibility test to identify its suitability for proceeding with data warehousing project or to proceed for improvement of the implemented operational system first. We also proposed an integrated approach of solution development where requirements of both operational system and data warehousing solutions could be considered together to achieve quality in both type of solutions and avoiding effect of poor operational system in data warehousing solution. We implemented data warehousing solution for one of the key subject areas of NRHM project and executed some useful queries over the developed solution and found that all the queries executed satisfactorily and produced results as per expectations of the business managers. Our implementation worked as a solution prototype and produced an early feedback of the users to proceed further with more implementation.

Declaration: The sole purpose of our work is analysis of the technical and managerial aspects of data warehousing. Material referred, names, terminology, and contents referred from concerned external materials are properties of respective owners. We have used them purely for the academic enhancement and not for any other purpose or intentions. Our objective is not to criticize any system. We have tried to investigate and suggest possible improvements.

References
[1]. Rami Rifaieh, Nabila Aïcha Benharkat, "Query-based Data Warehousing Tool", DOLAP'02, November 8, 2002, McLean, Virginia, USA. Copyright 2002 ACM 1-58113-590-4/02/0011
[2]. STAMFORD, CONN February 24, 2005
[3]. Jukic, N., (2006), "Modeling Strategies And Alternatives For Data Warehousing Projects.", Communications of The ACM April
[4]. 2006/Vol. 49, No. 4, p-83-90
[5]. http://www.embarcadero.com,Tech Notes, "Why Data Warehouse Projects Fail", January 2008.
[6]. Predicts2012,"business_intelligence_still_subject_to_non-technical_challenges"
http://www.saiindia.gov.in/english/home/our_products/Audit_Report/Government_Wise/union_audit/recent_reports/union_perform ance/2009_2010/Civil_%20Performance_Audits/Report_no_8/chap_9.pdf
[7]. Chris JenningsInsights, "Five Key Causes on Why Data Warehouse Projects Fail", December 22, 2014
[8]. Deepak Asrani, Renu Jain, & Usha Saxena, "Limitations of Operational System Design and Requirements of a Successful Data Warehousing Solution: A Gap Analysis for NRHM", 4th International Conference on Emerging Trends in Engineering and Technology College Of Engineering ,Teerthanker Mahaveer University, 24-25 Apr, 2015