Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com

BI Portal for Healthcare Informatics

Urvisha Patel[1], Kedar Sarang[2], Chetan Prabhu[3], Seema Yadav[4]

[1,2,3]Student, [4]Assistant Professor, Department of Information Technology, K.J Somaiya Institute of Engineering and Information Technology, University of Mumbai, Maharashtra, India

Abstract: **Business intelligence (BI) and healthcare analytics are emerging technologies that provide analytical capability to help healthcare industry improve service quality, reduce cost, and manage risks. However, such component on analytical healthcare data processing is largely missed from current healthcare information technology or health informatics curricula. The paper refers to the use of services like SSIS, SSAS, SSRS to provide integration, analysis and reporting of humongous data, respectively. The project conducts a analysis on how business intelligence can be incorporated in the field of healthcare informatics for betterment, enhancement in the field .A general framework and exemplar implementation strategy is presented.**

Keywords: **BI, BI Portal, Business Intelligence Development Studio (BIDS), ETL, Healthcare Analytics, Internet Information Services (IIS), Sql Server Integration Service (SSIS), Sql Server Analysis Service (SSAS), Sql Sever Reporting Service (SSRS).**

I. INTRODUCTION

The worldwide need for comprehensive and improved patient care and reduced healthcare costs has increased the emphasis and importance of healthcare data management and analysis. Healthcare depends on diverse and unique sources of data to support the diagnosis, treatment and prevention of disease, illness, injury and other physical and mental impairments and damage to human life, as well as the operation of healthcare service providers and organizations.

Healthcare data has some unique features which include diversified sources of data, convolution, and entanglement of different datasets, level of regulation and management, the potential impact on a patient's health and life. Healthcare data includes patient information, clinical data, financial data, medical knowledge, and operational data. For example, the data in a hospital or clinic is in various and differing formats and comes from disparate and contradictory sources including clinical and operational data. As the result, such data is difficult to analyze and often fail to convey its importance, its value to the laymen non-technical user thus deteriorating its delivery value and perception. Another major challenge to healthcare IT today is data growth is volatile and dynamic, such as digital imaging and electronic health records in general with the data in use being colossal and gigantic.

This is where business intelligence (BI) systems come in to help integrate, analyze report and manage data to, turn data into in meaningful information, and therefore improve patient care, reduce costs, and optimize business usage and understanding of the data. BI is believed to be the key for enhancing healthcare quality with less cost thereby increasing the overall efficiency of the entire process.

Demand for BI applications for healthcare grows incessantly as the volume of data and the desire to learn from the data keeps on expending. The demand for data management and analysis expertise in healthcare is increasing at phenomenally quick rate. So this project gives a general framework and presents several exemplar implementation strategies as well as conducts an elementary and exploratory analysis on how healthcare business intelligence can be incorporated into a HIT program thus resulting in perfect amalgamation of everything that is required to create a revolutionary impact in the field of healthcare informatics.

Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com

II. BACKGROUND

Business intelligence is a profound term to describe a group of methods, processes, architectures, applications, and technologies that transform quintessential, raw data into meaningful and useful information to support business operation and growth.

Lack of information or improper utilization of the information may result in bad decision which is not desirable to business prospects. BI could be a convenient tool in the hands of management which could help the management in making better decisions by providing on the fly analysis, generation of reports and by providing predictive analysis of information available etc.

Intelligent exploration, integration, aggregation and a multidimensional analysis of data originating from various information resources are some of the quintessential task that are to be performed by BI. Amalgamation of data from internal information systems of an organization and intergradations of data coming from the particular environment is done by systems of BI standard. For e.g. statistics, financial and investment portals and miscellaneous databases. Such systems are meant to provide adequate and reliable up-to-date information on different aspects of enterprise activities. Implementation of BI systems leads to ability to access, use and share data and information in an efficient and relevant way helps improve business performance and overall business growth.

III. LITERATURE REVIEW

Existing Systems	Proposed System
1. Manual entry of data.	1. Automatic abstraction of data.
2. Non availability of compact data.	2. Availability of compact and compressed data.
3. Difficulty in maintenance of humongous data.	3. Simple to maintain Humongous data.
4. Manual retrieval of data.	4. Collection of data dynamically.
5. May lead to inconsistency of data.	5. Inconsistency can be removed.
6. Unable to handle and rectify such problematic issues[8].	6. Adept handling or rectification of problematic issues like deduplication or data lineage issues.
7. Complex issues like nonconforming dimensions cause a lot of problems in generating final reports.	7. Non conforming dimensions not a problem at all.
8. Slowly changing dimensions cannot be handled quickly .	8. Rapid handling of advanced modelling issue like slowly changing dimension .
9. Manipulating and dealing with multiple hierarchies is a issue.	9. Multiple hierarchies is not an issue at all .

IV. ARCHITECTURE

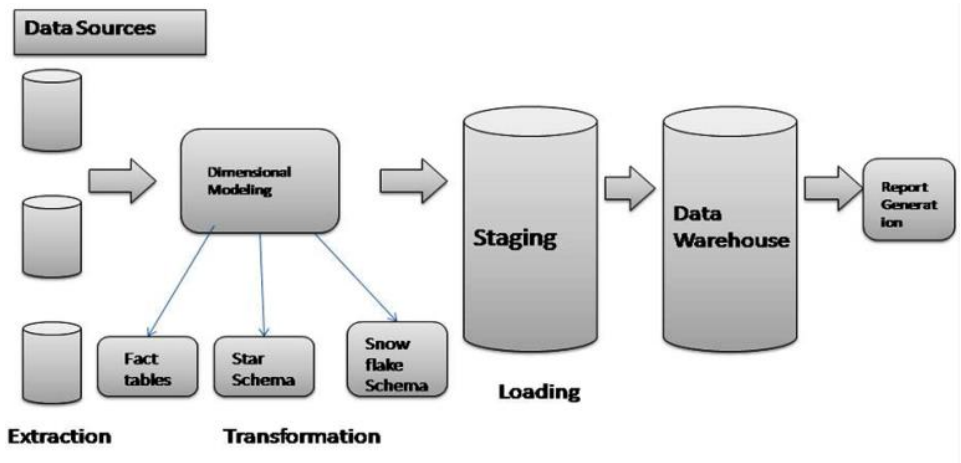


Fig. 1 Block Diagram

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com

Turning the data into consumable information that can facilitate decisions is not a luxury but a necessity to stay competitive in the business [8]. The term Business Intelligence (BI) incorporates the concept of deriving useful information from the data in an organization. The architecture in Fig 2 represents the same procedure.

A. Data Source:

One of the challenges when working with data in a BI system is that it typically originates in many different data storage systems. Extracting data from those different sources and merging the data into a single, consistent dataset is challenging. When working with diverse data sources, the following issues may be considered:-Different source environments with different systems including different platforms and operating systems, Source data may originate in many different types of database system, including Oracle, DB2, SQL Server or others.

1) Extraction: The first part of an ETL process involves extracting the data from the source systems. Most data warehousing projects consolidate data from different source systems [5][11]. Common data source formats are relational databases and flat files. Once the source data is studied, extracting meaningful data is the next big challenge with data integration.

B. Dimensional Modeling:

Dimensional modeling is the design concept to build the data warehouse. The dimensional data model provides a method for making databases simple and understandable [3].

1) Dimensional Table: It is a business entity of the source system. There can be multiple normalized table represent one single business entity on the source system [3].

Example: Customer Dimension, Product Dimension.

2) Fact Table: A central table in a data warehouse schema that contains numerical measures and keys relating facts to dimension tables [3]. There are basically two types structures that are popularly used while designing the physical dimensional modeling, they are Star Schema and Snow flake Schema.

3) Star Schema: The star schema has a center, represented by a fact table, and the points of the star, represented by the dimension tables. From a technical perspective, the advantage of a star schema is that joins between the dimensions and the fact tables are simple, performance, ability to slicing and easy understanding of data [3].

4) Snowflake Schema: A variation of a star schema is called a snowflake. In a snowflake, the dimension tables are normalized. From a reporting perspective, there is no logical difference because all the tables of a snowflake dimension can be logically joined into a single star dimension. From a performance perspective, the snowflake may result in slower queries because of the additional joins required [3].

5) Transformation: The transformation process allows us to consolidate, cleanse, and integrate data. We can semantically synchronize data from heterogeneous sources [5][11].

A transformation converts the fields of the source into the format of the target.

6) Loading: Data will be moved to the center of the target data warehouse table, it is usually the last step in the process of ETL [5][11]. As to the best way to load data, the implementation depends on the type of operation and the quantity of data.

C. Staging:

Data extracted, transformed & loaded from source to destination may require a temporary storage for various reasons such as addressing failures, reducing load on the source system, data cleansing, auditing purposes etc. Staging is therefore used for temporary storage of the data.

D. Data Warehouse:

Data warehouse is the core of any solid BI solution. Data Warehouse can be defined as a repository for keeping data in a subject oriented, integrated, time variant and nonvolatile manner that facilitates decision support [12].

Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com

E. Report Generation:

Finally the last part report generation. Reporting is essential for assessment and decision making. Healthcare industry needs a variety of reports in enterprise intelligence and medical intelligence to assess patient outcomes and quality of care, understand the potential impacts of many regulatory proposals. The data which is required is already stored in the data Warehouse & now it is ready for the generation of reports which could help the organization for the analysis of required data. The reports can be generated in different formats.

V. PROJECT METHODOLOGY

1. Design the data warehouse first through using SQL Server Management Studio. In particular, design the tables (dimensions and fact) that are needed as part of the Data warehouse [5]. Whenever the project is executed business intelligence development studio automatically fetches the data from the dataset & stored it in the required table in Sql server management studio that we have created so no manual entry of data is required. The tables or the entities taken for the healthcare are needed to be mapped with each other by their keys ,so as the relevant data and queries can be formed and their results could be obtained. The dataset could be of any format.

2. In Fig. 2 a sample table name patient is being created , the attributes given to it are the same as that of the dataset's patient table present in the excel format. As earlier mention that no manual entry of data is required after creation of the table the tuple values for all the attributes are filled in as shown in the Fig. 3

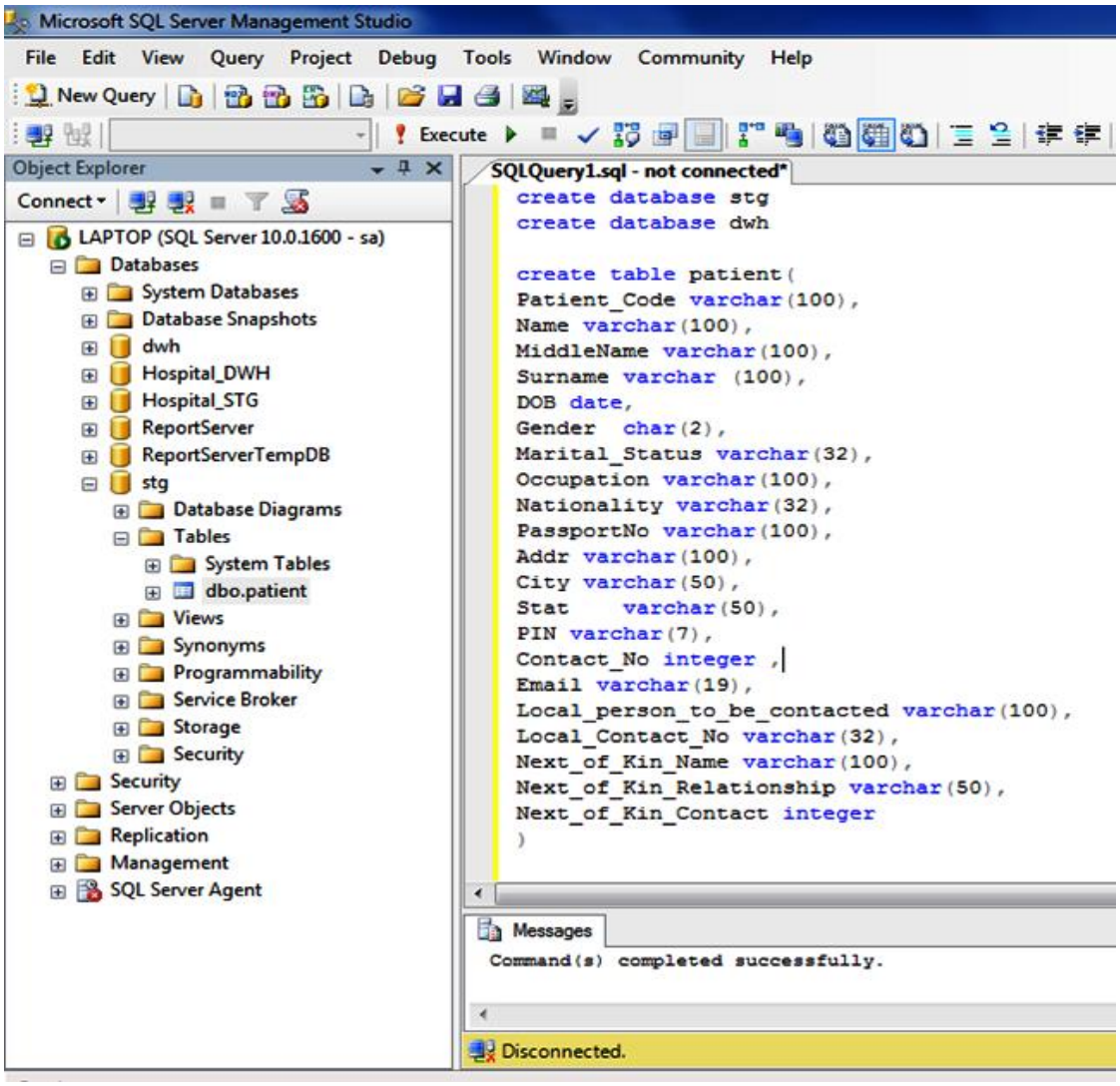
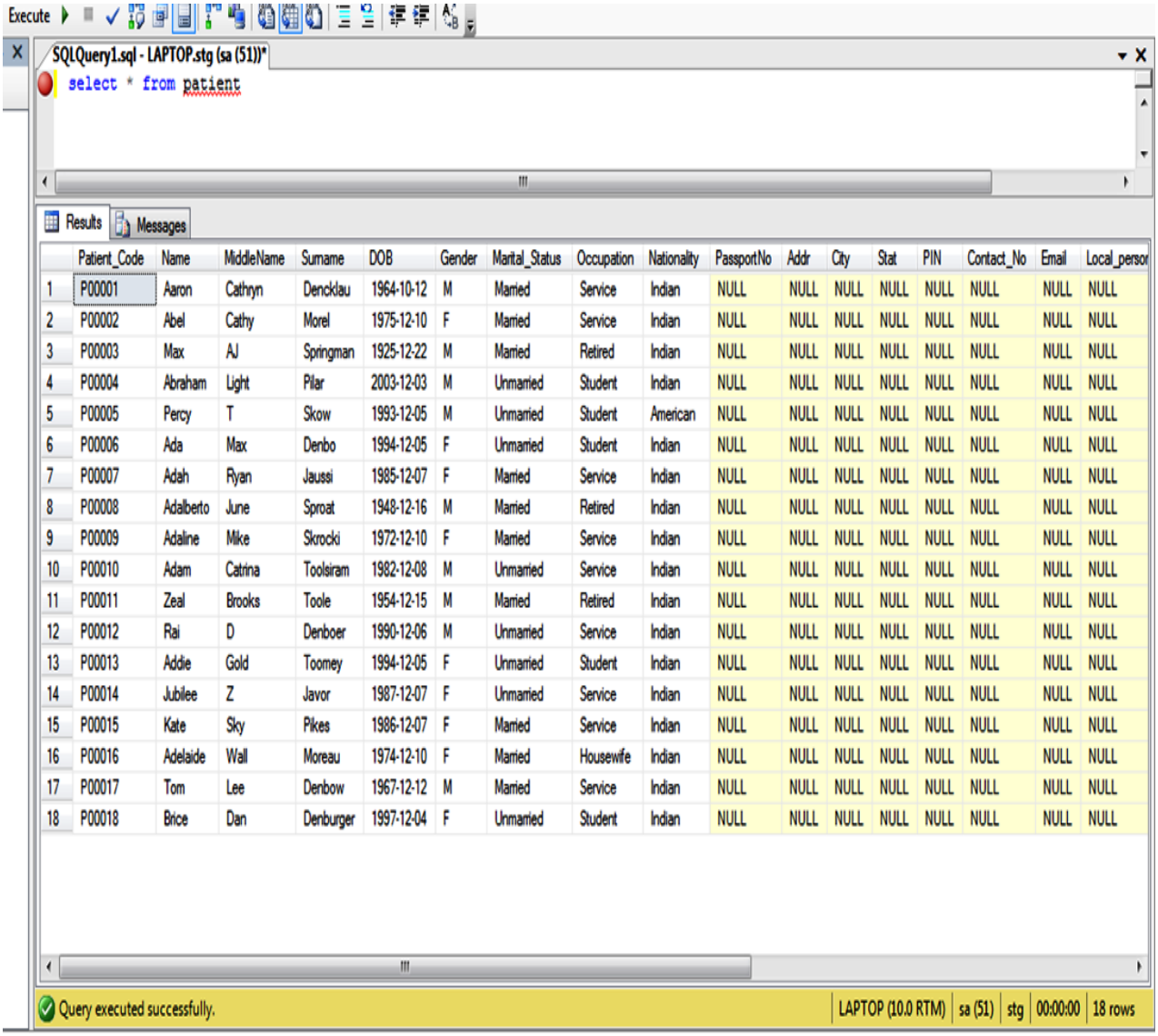


Fig. 2 creating the table

Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com



	Patient_Code	Name	MiddleName	Surname	DOB	Gender	Marital_Status	Occupation	Nationality	PassportNo	Addr	City	Stat	PIN	Contact_No	Email	Local_perso
1	P00001	Aaron	Cathryn	Dencklau	1964-10-12	M	Married	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2	P00002	Abel	Cathy	Morel	1975-12-10	F	Married	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3	P00003	Max	AJ	Springman	1925-12-22	M	Married	Retired	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	P00004	Abraham	Light	Pilar	2003-12-03	M	Unmarried	Student	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	P00005	Percy	T	Skow	1993-12-05	M	Unmarried	Student	American	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	P00006	Ada	Max	Denbo	1994-12-05	F	Unmarried	Student	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
7	P00007	Adah	Ryan	Jaussi	1985-12-07	F	Married	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
8	P00008	Adalberto	June	Sproat	1948-12-16	M	Married	Retired	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
9	P00009	Adaline	Mike	Skrocki	1972-12-10	F	Married	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
10	P00010	Adam	Catrina	Toolsiram	1982-12-08	M	Unmarried	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
11	P00011	Zeal	Brooks	Toole	1954-12-15	M	Married	Retired	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
12	P00012	Rai	D	Denboer	1990-12-06	M	Unmarried	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
13	P00013	Addie	Gold	Toomey	1994-12-05	F	Unmarried	Student	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
14	P00014	Jubilee	Z	Javor	1987-12-07	F	Unmarried	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
15	P00015	Kate	Sky	Pikes	1986-12-07	F	Married	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
16	P00016	Adelaide	Wall	Moreau	1974-12-10	F	Married	Housewife	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
17	P00017	Tom	Lee	Denbow	1967-12-12	M	Married	Service	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
18	P00018	Brice	Dan	Denburger	1997-12-04	F	Unmarried	Student	Indian	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Fig. 3 Patient table

3. Next step would be of the ETL process that is done by Sql Server Business intelligence development studio, which contains a lot of data integration components that can be drag & drop, simultaneously an ETL database is created in SQL Server management studio as shown in the Fig. 4.

i) Various data flow sources are available like excel source, OLE DB source, Flat File Source, Raw File Source that are present in Business intelligence development studio so based on the type of data source to be extracted the component is selected. This is referred as the extraction process in ETL.

ii) Next is of transformation, similar to extraction components there are also various data flow transformation components available like for example that of derived column shown in Fig. 5. Derived column can be used for applying the rules. For example if the patient table present in the data source consist of an attribute name gender and whose values are given as 'M' or 'F' and according to rules of data warehouse if it wants the value to be as 'MALE' or 'FEMALE' then at that time derived column component could be used for transformation.

iii) Next is loading for which data flow destination components are available .We can select any of the destination but generally OlE DB destination is preferred as it is a standard bridge for many data sources. By this step we now have all the records of datasources in their respective tables available in our ETL database as per the rules.

Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com

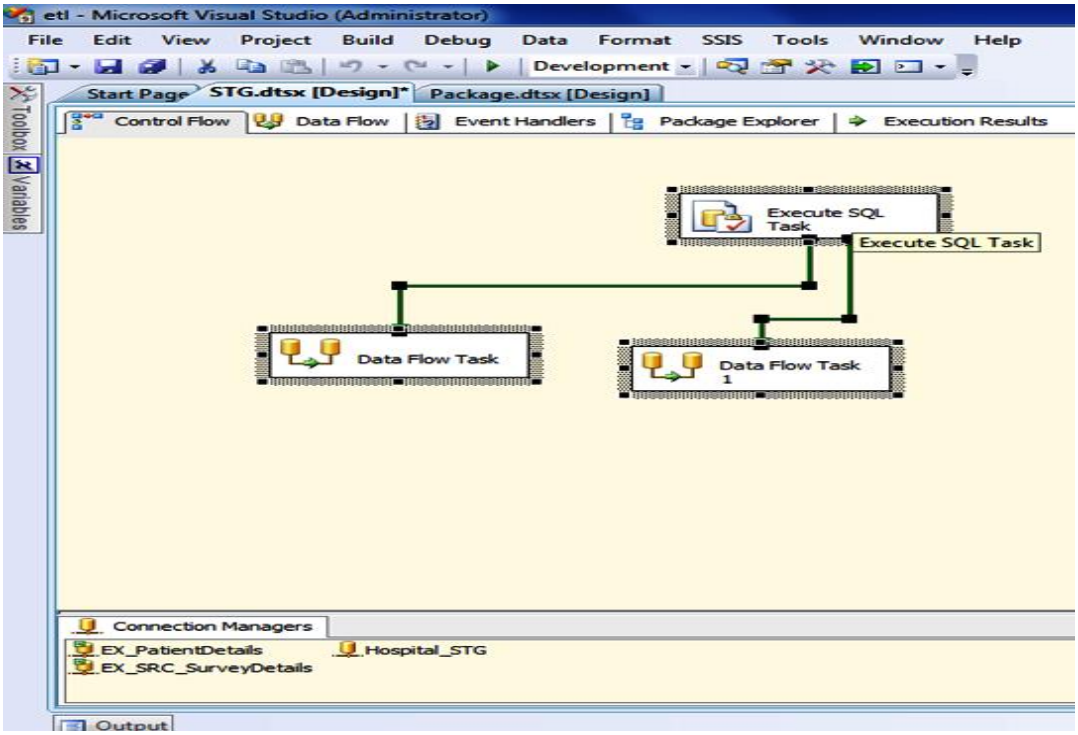


Fig. 4 Building the ETL using BIDS

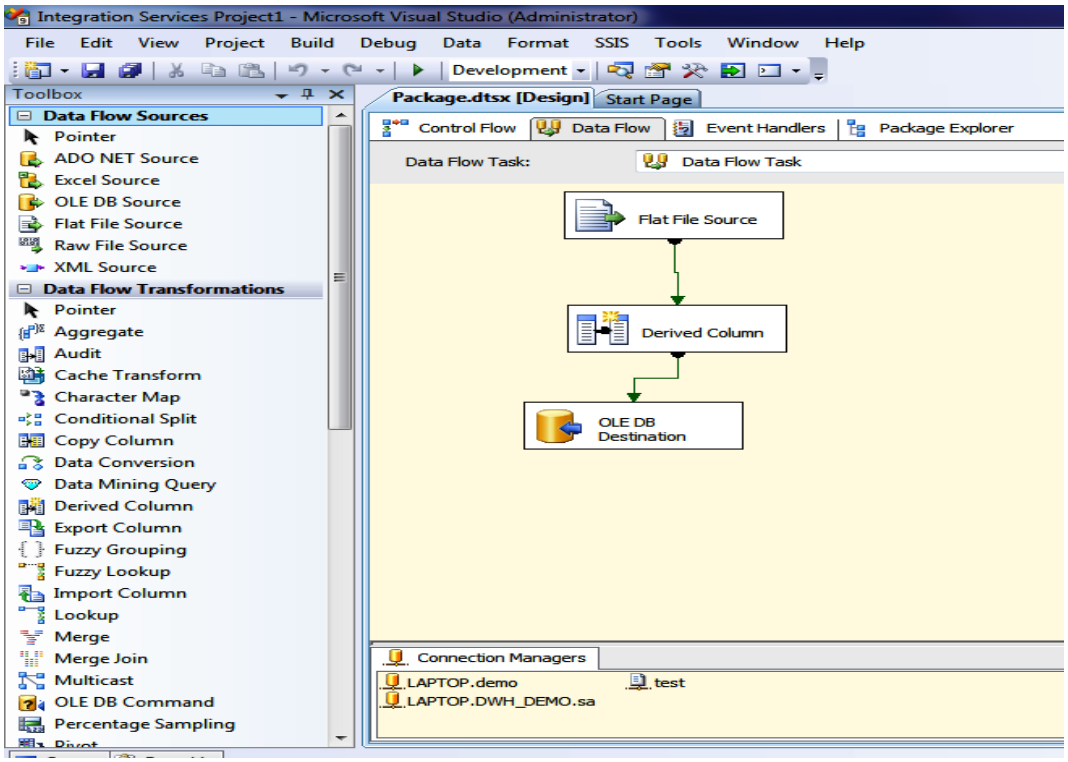


Fig. 5 Data Flow Transformation

4. Since BI deals with colossal historical data, there has to be a quick way to obtain the results. Directly executing the queries on the tables may lead to increase in query execution which is not desirable in BI. So in order to increase query performance data warehousing concept is used. In data warehouse each table is decomposed into number of table in order to increase the query optimization as shown in the Fig. 6.The reporting process is actually linked with the data warehouse. So whenever the reports are generated the queries are fired on these data warehouse.

Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com

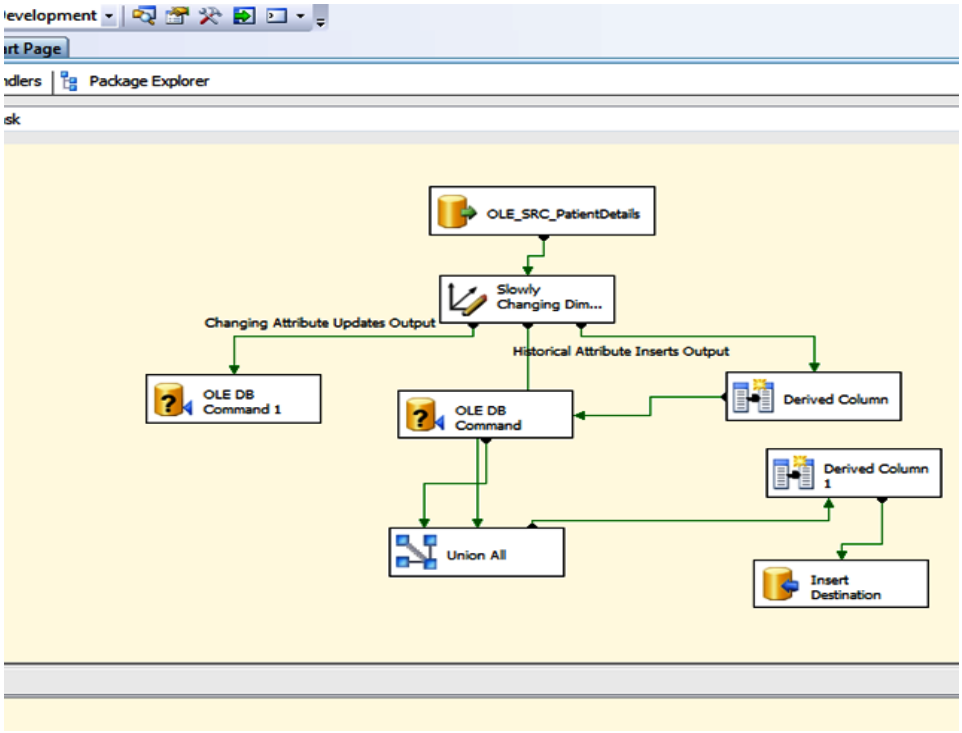


Fig. 6 Decomposing of tables in DW

5. Before initiating the process to display the reports on the browser there is a need to start the Reporting Service Configuration Manager software. So by this we are linking the entire reporting project with Internet Information Services (IIS), which will help us to display the reports. The report can be generated in any format and its output will be shown in internet explorer through local host.

Fig. 7 is one such example of a report that can be made. Also we can linked the report that is generated to any other website.

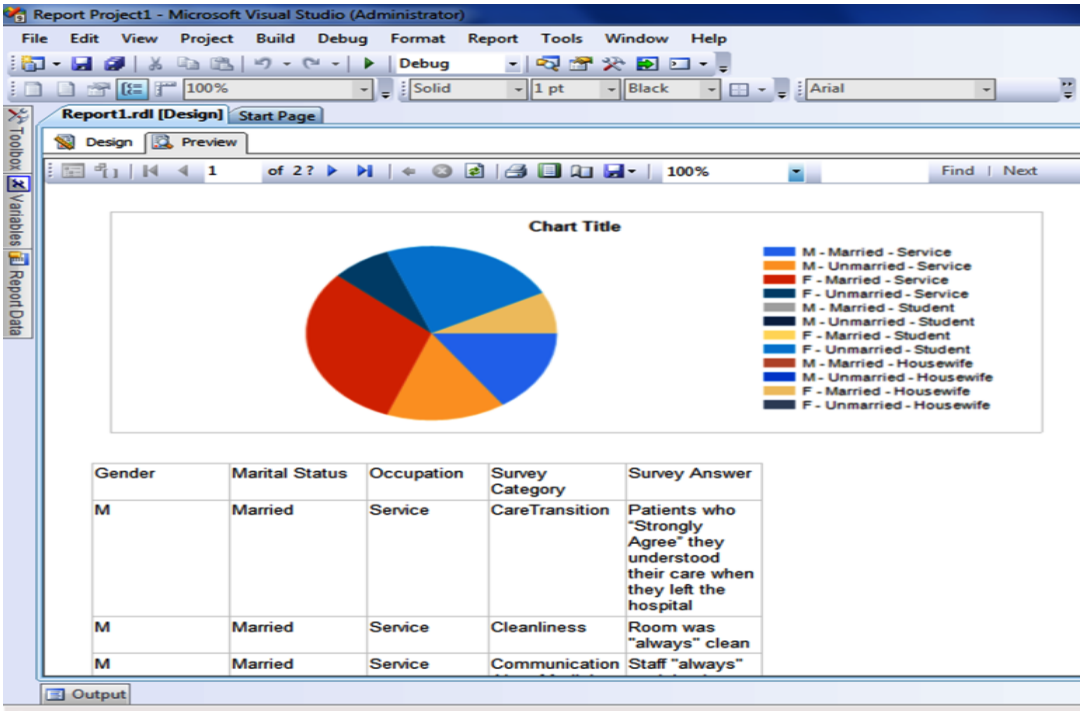


Fig. 7 Sample report

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 3, Issue 1, pp: (38-45), Month: January-April 2016, Available at: www.noveltyjournals.com

VI. CONCLUSION

The reporting, analysis, and interpretation of business data is of umpteen concerns for any organization in order to guarantee its competitive edge, optimizing processes, and enabling it to react quickly and adapt to varied changes. BI Portal provides flexible reporting, analysis, and planning tools to support Healthcare in evaluating and interpreting data which would result in increasing the efficiency of managing any particular healthcare organization.

In today's world, where medical technology is transcending the pinnacles of success for improving public health, it is very essential that people get access to these ever improving medical help which is only possible through proper implementation and usage of healthcare informatics. Here comes the concept of business intelligence which is cornerstone of the success of implementation and usage of healthcare informatics, as our project amalgamates business intelligence technology and healthcare informatics for the betterment of all the people concerned with healthcare system including its providers as well as customers, this project is of highly social relevance.

REFERENCES

[1] Guangzhi Zheng, Chi Zhang and Lei Li "Bringing Business Intelligence to Healthcare Informatics Curriculum: A Preliminary Investigation" pp.205-210 ,2014.

[2] Zhijun Ren , "Constructing a Business Intelligence Solution with Microsoft SQL Server 2005" ,2010.

[3] Ralph Kimball and Margy Ross , "Kimball Dimensional Modeling Techniques" ,2013.

[4] Ton A.M. Spil, Robert A. Stegwee, Christian J.A. Teitink , "Business Intelligence in Healthcare Organizations" Proceedings of the 35th Hawaii International Conference on System Sciences – 2002.

[5] Osama T. Ali, Ali Bou Nassif and Luiz Fernando Capretz , "Business Intelligence Solutions in Healthcare A Case Study: Transforming OLTP system to BI Solution" pp. 209-214,2013.

[6] Maria Antonina Mach, Abdel-Badeeh M. Salem , "Intelligent Techniques for Business Intelligence in Healthcare" pp 545-550 , 2010 10th International Conference on Intelligent Systems Design and Applications.

[7] G R Gangadharan , "Business Intelligence Systems: Design and Implementation Strategies" pp 140-144, 2dh Int. Conf. Information Technology Interfaces IT1 2004.

[8] Celina M. Olszak, Kornelia Batko ," The Use of Business Intelligence Systems in Healthcare Organizations in Poland " pp. 969–976, Proceedings of the Federated Conference on Computer Science and Information Systems,2012.

[9] Mana Azarm, Fatemeh Nargesian, Liam Peyton ," Managing and Mapping Data Lineage for Business Intelligence and Analytics Applications in Health Care " pp-120-126 ,2011.

[10] Patti Brooks, Omar El-Gayar, Surendra Sarnikar ,"Towards a Business Intelligence Maturity Model for Healthcare ", pp-3807-3816, 2013 46th Hawaii International Conference on System Sciences.

[11] TangJun,Cui Kai,Feng Yu,Tong Gang, "The Research & Application of ETL tool in Business Intelligence Project" , pp-620-623,2009 International Forum on Information Technology and Applications.

[12] Zhijun Ren, "Building Business Intelligence Application with SAP BI " ,2009.