International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Data Warehouse Testing: A Step Towards using Business Intelligence in Banking

Sonali Mathur, S.L. Gupta, Payal Pahwa

Abstract: ***In the modern era, technology plays a vital role by contributing in the development of business by providing various tools and techniques to enhance the business decision-making process. Data warehouse is an important entity that contributes to the decision-making process, which can be seen in the literature available over the years. Data warehouse provides the basis for quality analysis of available data by deriving accurate information from data. Like many other industries, banking sector is also facing challenges due to various reasons like large over-dues, non-performing assets, changing customer demographics, matching customer expectation levels, increased competition from financial technological companies and banking competitors, etc. Thus data warehouse system serves to be the best solution for the banks to overcome challenges as data warehouse system integrates all the data at one place and provides a consolidated view of the past transactions which can be used for report generation and performing analytical analysis in order to help the management to maximize business performance. In addition to making strategic decisions, data warehouse also assist in helping the banks to improve customer retention, optimize discounting, market segmentation, business performance, customer deposits, etc. Thus, a data warehouse system provides a solution to all data management problem and generate patterns and reports for analytical end users for enhancing decision making processes. In this research paper, we shall be representing a bank model that focuses on the loan department of a bank data warehouse and shall be explaining how business intelligence plays a role in improving the loan analysis for the banks. Loan analysis may include summarizing the classified loans, analyzing loans within and exceeding the threshold limits of Loan-To-Value (LTV) ratio, analysis of loan rejections and other analysis pertaining to loans.***

Keywords : Data Warehouse, Data Warehouse Testing, Banking Industry, Business Intelligence.

I. INTRODUCTION

Analytics focusses on turning raw data into meaningful information so that it can be used for making better decisions. Analytics is based on the statistical application, computer programming and operations research in order to gain meaningful information from the data. This field is especially useful in areas where lot of information exists. Banking industry is the area where lot of transactions are executed on daily basis, thus banks have become a source of lot of information. While basic reporting and descriptive analytics are mandatory services being carried out on regular basis in banks, advanced predictive and prescriptive analytics are now starting to generate powerful insights resulting in significant business impact. Structuring and recording the data is useless until the data is not utilized for generation of meaningful information. Thus, for banks the data can be used for maintaining security, detecting frauds, managing risks, customer analysis, customer investment pattern analysis, market segmentation, product customisation, etc. For a business to grow they constantly required a database environment with high flexibility, better adaptability and good support for making decisions. One of the best solutions that can be thought of is the use of data warehouse technology for handling large databases and gaining insight into the information collected over the years and assist the higher management in making business decisions for the business growth [24]. According to Inmon [1] data warehouse systems are defined as a central repository which contains data collected from various sources, integrates and transform the data through the Extract, Transform and Load process (ETL). The users can use the appropriate data analysing tools to store and analyse the data to generate reports and perform analysis. Data warehouse applications helps the analysts to access, aggregate and analyse huge volume of historical data to forecast and achieve the expected result for making smart decisions [2]. The major benefits related to the data warehouse applications include saving of time, availability of good quality information which may help the management to take tough business decisions in order to improve the processes and help in achieving strategic business objectives [3, 4]. In this research paper we have discussed the impact of data warehouse testing on the banking environment. The paper has been organized in the following way – Section 2 talks about the uses of Business Intelligence in Banking sector. Section 3 describes the contribution of various researchers in this area. Section 4 briefly explains what is data ware house testing along with Section 5 describing the testing goals and Section 6 represents the data warehouse testing framework on which the bank model has been made. Section 7 and 8 discusses the various testing techniques used in data warehouse testing and the its implementation through a bank model. Section 9 and 10 describe the results and conclusions of the research work.

Manuscript published on November 30, 2019.
* Correspondence Author
Sonali Mathur*, Research Scholar, BIT (Mesra, Ranchi), India.
Prof. (Dr.) S. L Gupta, Dean, Waljat College of Applied Sciences, Sultanate of Oman.
Prof. (Dr.) Payal Pahwa, Principal, BPIT, Rohini, Delhi.

II. ROLE OF BUSINESS INTELLIGENCE (BI) IN BANKING

Business intelligence nowadays is considered as a separate discipline comprising of certain elements of information technology, strategy, managerial processes, organizational analysis and marketing. It helps in resolving management problems by integrating, analyzing, disseminating and acting based on the business information in order to make the best business decisions [5].

Business intelligence system offers technological platforms and knowledge for its implementation thus, not existing as a final product. Banks are among the leaders in adopting new technologies and knowledge. Thus, special databases called as data warehouses are created to meet the needs of these BI systems, where data is organized in such a way that it convenient to conduct analytical processes on large volume of data sets. These data sets are structured specially for report and analysis. Data warehouses and OnLine Analytical Processing (OLAP) form the informational basis for applying business intelligence. Business intelligence takes care of many areas of the banking business, few of them being

- Analytical Customer Relationship Management
- Bank Performance Management
- Asset and Liability Management
- Enterprise Risk Management
- Compliance

Business intelligence is being used overall in every industry just to know from what happened to why it happened and what corrective solutions are required. The reasons why business intelligence can be helpful to the banking industry are:

1. Helps in getting faster answers to our business questions – Business Intelligence makes the visualization of data easier through reports and summary sheets, thus quick decisions can be made.

2. Elimination of guess work – Most of the executives rely on "best guess" and "gut feel" decisions to predict future decisions. Thus, with business intelligence, which provides more accurate historical data and helps in forecasting and trend analysis, guess estimates can be eliminated.

3. Get an insight into customer behaviour – Business intelligence enables the organization to gain visibility into customers behaviour by their historical trend analysis, thus can help in retaining their valuable customers.

4. Improve efficiency – Reporting analysis and visual representations help the executives to see where the business had been, where it is now and where it is going to be.

5. Streamline of business operations – Business intelligence helps in gaining insight into the business performance, thus organizations can easily see where they need to make changes to streamline their operations.

Figure 1 represents the business intelligence cycle which is a continuous cycle of data analysis, insight, action and feedback. There are three success factors that are to be focused if business analytics solutions are to be implemented in the banking sector. They are:

1. Data used for creating business insights.

2. Decisive actions taken using the business insights.

3. Results from business decisions are used to improve business performance.

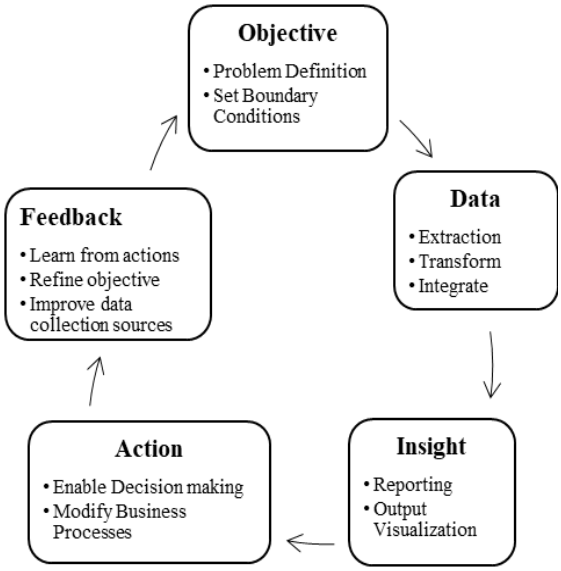


Figure 1. Business Intelligence Cycle

III. RELATED STUDIES

The idea of data warehouse was way back introduced by [6] where the author had suggested to construct read-only databases that stores the historical data and can be used for querying, analysis and search activities which are then used for decision making purposes. The data warehouse systems as defined by [1], mainly have four characteristics, "subject-oriented", "integrated", "time-variant" and "non-volatile". "Subject-oriented" attribute refers to considering a database related to a certain topic as we are considering the loan department of a banking enterprise. "Integrated" implies integrating the data from various sources, as for the data source, we have integrated the data from various branches of a bank. "Time-variant" focuses on the time interval, as we have considered historical data of last four years to be queried and used for analysis purposes. "Non-volatile" refers to the integrity of the historical data which can be used for analysis, prediction and discovery of patterns to help analysts to make better decisions. No modifications are done on the data once it is loaded into the data warehouse in order to preserve its quality. In the early 1980s, relational database management systems were used as commercial products due to their simplicity along with the query capabilities provided by Structured Query Language (SQL), enhanced and supported the interest in the business intelligence and decision support systems [7]. Thus, data warehouse had been established since then in many industries like telecommunication and logistics, manufacturing and production (for order management and customer support) to health and financial services (for customer retention, risk analysis, etc.) [8]. Data warehouse working on historical data are the most suitable solutions which can be used by decision makers [9] as they are not working on the real time data environment. Therefore, data warehouse system can be seen as an informational environment [10], that (i)

gives a complete and integrated view of an enterprise; (ii) makes the enterprise's historical and current information easily available for strategic decision making;

(iii) helps in making decision support transactions possible without affecting the operational systems; (iv) gives consistent information to the organization; (v) shows a flexible and interactive source of generating strategic information. Data warehouse can thus be called as a user centric environment, where the users can generate strategic information based on the available data in order to improve business decisions. The data warehouse can be classified into three components: (i) Data Collection, (ii) Data Storage (iii) Data Analysis.

(i) Data collection – The data is collected at the back end of the data warehousing system and consists of systems that have interfaces to load data into the data warehouse from operational systems [11]. Data from operational environment based on OnLine Transaction Processing (OLTP) systems which may consist of databases such as Oracle, MySQL, SQL Server, DB2,etc are loaded into the data warehouse and subjected to Extract-Transform-Load (ETL) operations. Based on the data warehouse design, the data is extracted external or internal sources, transformed into a specific format and loaded into the data warehouse. The process of transformation improves the quality of the data by removing inconsistencies, in accurate data, missing values duplicate values and NULL values in data [12].

(ii) Data Storage – The data stored in the data warehouse can be used for two purposes – either to get fast response of the queries given by the user or the data is used by the analysts to perform analysis on the given set of data. Data is stored in the form on dimension table and fact tables. For simplifying usage and analysis of data small units of organization called as "data mart" are considered such as we have considered the loan department of a banking organization.

(iii) Data Analysis – This component is termed as the front end of the data warehouse, where the business applications use the data stored in the data warehouse for performing analysis and generating reports. Variety of tools can be used for this purpose such as data analytical tools, data mining, machine learning, etc. OnLine Analytical processing (OLAP) and data mining are the two best techniques used for discovery of knowledge [13]. OLAP and data mining techniques are used for generating reports and interesting patterns that can be studied for improving business decisions. Huge amount of data is processed by decision support systems (DSS) queries, thus a simulator was proposed for optimizing the queries in order to increase the performance of the queries [23].

IV. DATA WAREHOUSE TESTING

Data is passed through various processes in a data warehouse which may bring about changes in the data before it finally reaches to the user in the form of a report or a graph. To verify that the data quality has not decreased during the transformation processes, data should be tested at every phase of development in order to preserve the quality of data. Verification and validation are two important components of testing. Verification ensures that the product being used meets the user's expectations. Validation on the other hand, is carried out from the specification through control, design, test case preparation, execution and evaluation of results. Thus, testing is one of the basic tool of validation [14]. For organizations to make accurate decisions, testing should be planned and executed efficiently in order to prevent erroneous data in the database which may be reflected in the reports generated to be studied by the higher management for decision making process. Testing is a process executed with the aim to detect errors as stated by [20] is shown in Fig 3.1. The difference between software testing and data warehouse testing has been explained and data warehouse testing is done on business logic [15]. Less focus on data warehouse teting has lead to failure of many data warehouses [25]. Software stratgies are adequate for use in data warehouse testing [26, 27, 28]. Good test cases are formed so that they are able to detect errors. The idea is to detect errors at minimum cost and time. Test cases are then executed using the data available to check whether the received output value is similar to the expected outcome value or not. If the values are similar then no error exist, otherwise error detected is removed. Before studying the testing techniques let us focus on the goals of testing explained in the next section.

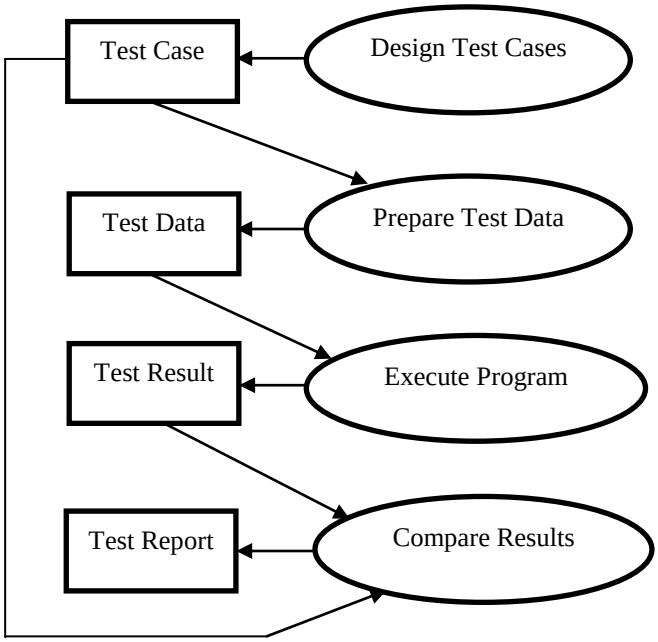


Figure 2. Testing Process [20]

V. TESTING GOALS

There is extra cost associated with finding errors in the later half of the development cycle. In data warehouse, incorrect data will lead to incorrect business decisions. The goal of testing a data warehouse application can be given as:

1. Data Completeness – It ensures that all data is loaded.

2. Data Transformation – It ensures all data is transformed according to business rules and specification.

3. Data Quality – It ensures that all inconsistencies in data are removed such as in accurate value, missing value, duplicate value, NULL values, etc.

4. Performance and Scalability – It ensures that data is loaded within timing constraints and queries are executed within expected time frame.

5. Integration Testing – It ensures that the ETL processes have executed properly.

6. User Acceptance Testing – It ensured that the solution meets the user expectations.

7. Regression Testing – It ensures that the existing functionality remains unchanged irrespective of the new code releases.

VI. DATA WAREHOUSE TESTING FRAMEWORK

The classical data warehouse testing framework consists of five components as shown in Figure 3, the data source, security in ETL, data warehouse, user interface for executives and reports. We have considered a bank model to work on this framework. The bank database is collected over last few years as the data source. The data source contains historical bank data for performing long term analysis. ETL processes (extract, transform and load) are executed on the source data in order to clean the data. The cleansed data is encrypted by using the Blow fish encryption algorithm in order to secure the data so that it cannot be accessed by unauthorized users. Then the transformed and encrypted data is loaded in the data warehouse. Test cases are generated and executed on the bank data warehouse to perform analysis of the data and generate reports which can be used by the higher management for decision making purposes.

for application domains like reporting, OLAP and information presentation more testing tools can be examined. Data warehouse testing methodologies have been proposed by many researchers as existing in the literature which can be seen in [21, 22]. Based on these methodologies, data warehouse testing techniques have been explained by many authors and have been compiled in [15] and it has been concluded that for front-end data warehouse testing, the best techniques which can be used are performance, regression, security and report testing. Keeping this in mind, we have tested a bank data warehouse model of loan department, considering few test cases which have been mentioned in the next section and evaluated the results generated. The test results generated shall help the higher management in making business decisions for increasing growth and profitability of the banking enterprise.

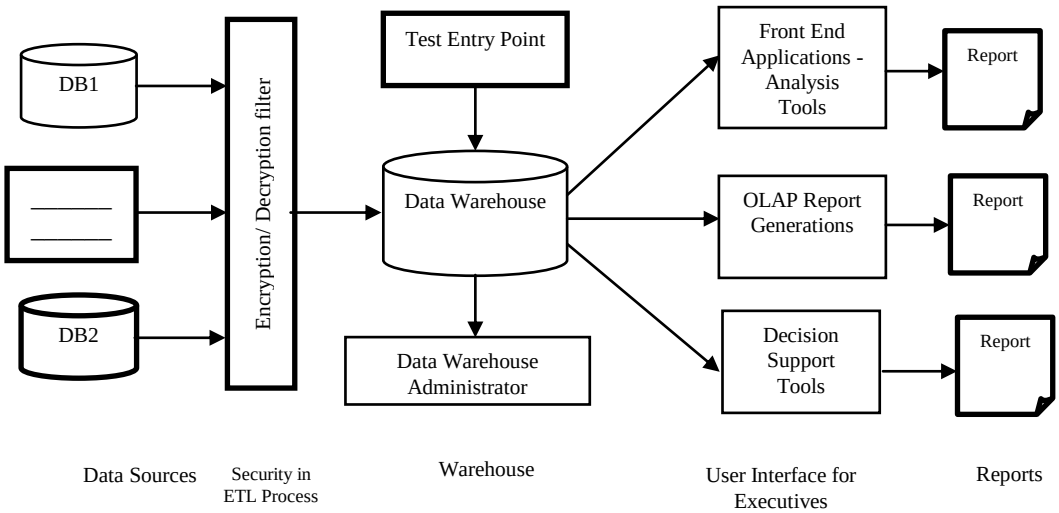


Figure 3. Data Warehouse Testing Framework

VII. DATA WAREHOUSE TESTING TECHNIQUES

Many organizations are nowadays depending on the data warehouse as data is available everywhere and the basic fact is that data warehouse contain data of high quality. To ensure that organizations make smart and accurate decisions, testing should be planned and executed efficiently so that the reports and statistics are generated from error free data thus supporting the management in making correct business decisions. Data warehouse testing is being explored by researchers in order to fetch meaningful information from the data. Some of the consulting firms and IT companies have promoted data warehouse testing such as [16, 17, 18]. Automated testing tools were also used for testing a data warehouse by [19], and it was concluded that DbFit, TSQLUnit and utPLSQL are the tools which can be used but

VIII. IMPLEMENTATION

A bank database of loan department for last four years individually was considered and template of one year is shown in Figure 4. The bank database does not consist records of current bank transactions, but it contains historical bank loan records as bank data warehouse is being considered. Data at the source is collected, it is cleaned and transformed using the Extract, Transform and Load (ETL) process into the format as desired consisting of required attributes of the bank database.

During the ETL process all the invalid, inaccurate, redundant and NULL data values have been removed and only the valid data is considered in order to maintain quality of data as shown in Figure 5 which represents the cleansed bank database.

1	age	job	marital	education	default	balance	housing	loan
2	30;unemployed;married;no;1787.0;no;no;19;oct;79;1;-1;0;unknown;NO							
3	33;services;married;no;4789.0;yes;yes;11;may;220;1;339;4;failure;YES							
4	35;management;single;no;1350.0;yes;no;16;apr;185;1;330;1;failure;YES							
5	30;management;married;no;1476.0;yes;yes;3;jun;199;4;-1;0;unknown;YES							
6	59;blue-collar;married;no;0.0;yes;no;5;may;226;1;-1;0;unknown;YES							
7	35;management;single;no;747.0;no;no;23;feb;141;2;176;3;failure;NO							
8	36;self-employed;married;no;307.0;yes;no;14;may;341;1;330;2;other;YES							
9	39;technician;married;no;147.0;yes;no;6;may;151;2;-1;0;unknown;NO							
10	41;entrepreneur;married;no;221.0;yes;no;14;may;57;2;-1;0;unknown;NO							
11	43;services;married;no;-88.0;yes;yes;17;apr;313;1;147;2;failure;YES							
12	39;services;married;no;9374.0;yes;no;20;may;273;1;-1;0;unknown;YES							
13	43;admin.;married;no;264.0;yes;no;17;apr;113;2;-1;0;unknown;NO							
14	36;technician;married;no;1109.0;no;no;13;aug;328;2;-1;0;unknown;YES							
15	20;student;single;no;502.0;no;no;30;apr;261;1;-1;0;unknown;YES							
16	31;blue-collar;married;no;360.0;yes;yes;29;jan;89;1;241;1;failure;NO							
17	40;management;married;no;194.0;no;yes;29;aug;189;2;-1;0;unknown;YES							
18	56;technician;married;no;4073.0;no;no;27;aug;239;5;-1;0;unknown;YES							
19	37;admin.;single;no;2317.0;yes;no;20;apr;114;1;152;2;failure;NO							
20	25;blue-collar;single;no;-221.0;yes;no;23;may;250;1;-1;0;unknown;YES							
21	31;services;married;no;132.0;no;no;7;jul;148;1;152;1;other;NO							
22	38;management;divorced;no;0.0;yes;no;18;nov;96;2;-1;0;unknown;NO							

Figure 4. Bank Database

Once, the data at the source is collected, it is cleaned and transformed using the Extract, Transform and Load (ETL) process into the format as desired consisting of required attributes of the bank database. The hypothetical data set of the customer is represented in Table 1. The dataset of data warehouse of the loan department of a bank consists of multidimensional database as data is integrated from various sources. On-Line Analytical Processing (OLAP) is a multidimensional data modelling technique which views data in the form of a data cube. The OLAP cube stores the database information and allows the analysis at different the conceptual levels. The OLAP cube consists of various parameters that are predefined by the users which can aid in data analysis as shown in Figure 5.

queries. The final reports are used by the management for analysis and decision making purposes.

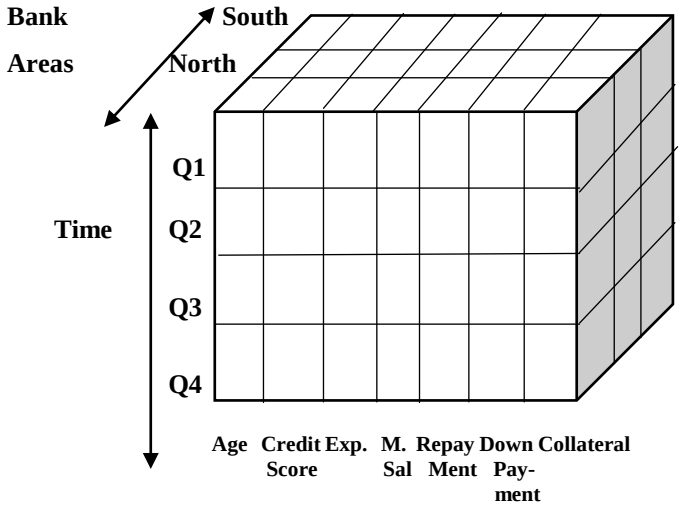


Figure 5. Data Cube for Loan Management Database

For the approval of the loan, certain parameters are define based on which the decision w.r.t the approval of loan is taken. The various parameters are:

1. **Age** – It is the key factor that evaluates how much premium is to be given to an individual and also the repayment period.
2. **Credit Score** – It is three digit number representing the numerical summary of the customer's credit history. It ranges from 300 – 800. It is also called as CIBIL Score.
3. **Experience** – It represents the working experience of the customer in an organization.
4. **Monthly Salary** – It is the total monthly salary that a customer generates per month and can be calculated as net income of the customer per financial year. It helps in determining the maximum loan value.

Table 1. Hypothetical Customer Dataset

Age	Job	Cust Id	Marital Status	Defaulter	Credit Score	Exp	M. Salary	Loan Amt	Repayment Years	Down Payment	Collateral	Quarter
29	Yes	1	No	No	338	4	68872	195020	7	16	Yes	Q1
59	Yes	2	Yes	Yes	685	4	18423	138177	4	42	Yes	Q1
25	No	3	Yes	No	479	2	57289	167004	5	21	No	Q3
58	Yes	4	No	No	510	1	60419	135507	5	28	Yes	Q3
51	No	5	No	Yes	603	4	77545	199932	7	38	Yes	Q4
62	No	6	No	Yes	702	1	48221	109863	6	49	Yes	Q1
45	No	7	Yes	No	362	1	43466	158901	7	49	Yes	Q4

Data cube store large banking data which are used by the administrator or management. OLAP is multi-dimensional data modelling techniques and to view the data is in the form of data cube.

For sanctioning of loans, the data of the customer, bank and the parameters required for loan approval are considered. OLAP operations describe the different levels of the system. In the case study hypothetical dataset for bank and customer is taken and implementation of the data cube is done through

5. **Loan Amount** – It is the amount for which the customer want to borrow the money from the bank.
6. **Years of repayment** – It is the time period during which the customer will repay the loan amount to the bank.
7. **Down Payment** – It is the amount that the customer is going to make the direct initial payment and on the remaining amount the customer wants the loan.

More the down payment, less is the monthly instalment for loan and less is the interest paid to the bank for the loan.

8. **Collateral** – It is an asset that the bank accepts as security for a loan. IF the borrower defaults on the loan payment, bank can seize the collateral and resell it to recover its loss.

The conditions on which the loan amount can be sanctioned is shown in the Table 2.

Table 2. Parameters for Loan Approval Process

Parameters	Conditions	Outcome
Age	The age should be between 25-60 years	Beyond 60 years of age, loan cannot be sanctioned.
Credit Score (CIBIL)	300 - 900	Below CIBIL score of 400, loan cannot be sanctioned.
Experience	1- 35 years	Below 4 years and above 35 years of experience, loan cannot be sanctioned
Monthly Salary	Rs. 10,000 – 50 Lac	Loan cannot be sanctioned for customer earning below Rs.15,000.
Loan Amount	Rs. 1 Lac – Rs. 1 crore	Loan amount eligibility is 60 times of net monthly salary.
Years of Repayment	2 years – 25 years	The shorter the repayment period, better is the eligibility for sanctioning of loan.
Down Payment	10% - 50%	Minimum 10% down payment is mandatory.
Collateral	Yes/No	Has any asset been provided as security to the bank.

Based on the above parameters and their range, decisions for loan approval can be done. The analysts can choose any parameter and test data will get loaded and test result is generated. Figure 6 represents test result generated after selecting age as one of the testing parameters for a specific year from the bank data warehouse. Similarly, for all the parameters the resultant data set satisfying the test case conditions is generated. Reports can be generated for the valid records to whom loan can be sanctioned based on defined parameters. Sample of a report generated after selecting any of the parameters is shown in Figure 7. Higher management can use these reports for analysis purposes in order to contribute in growth and profitability of the bank.

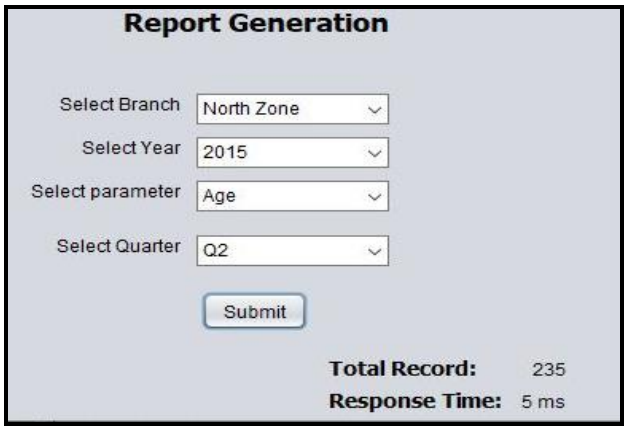
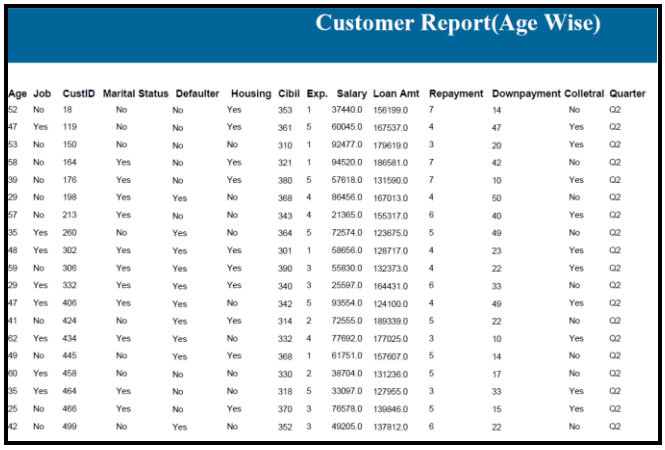


Figure 6. Sample of Test Result for Age Test case



Customer Report(Age Wise)

Age	Job	CustID	Marital Status	Defaulter	Housing	Cibil	Exp.	Salary	Loan Amt	Repayment	Downpayment	Colletral	Quarter
52	No	18	No	No	Yes	353	1	37440.0	156199.0	7	14	No	Q2
47	Yes	119	No	No	Yes	361	5	60045.0	167537.0	4	47	Yes	Q2
63	No	150	No	No	No	310	1	92477.0	179619.0	3	20	Yes	Q2
58	No	164	Yes	No	Yes	321	1	94520.0	186581.0	7	42	No	Q2
39	No	176	Yes	No	Yes	380	5	57618.0	131590.0	7	10	Yes	Q2
29	No	198	Yes	Yes	No	368	4	86456.0	167013.0	4	50	No	Q2
67	No	213	Yes	No	No	343	4	21365.0	155317.0	6	40	Yes	Q2
35	Yes	260	No	Yes	No	364	5	72574.0	123675.0	5	49	No	Q2
48	Yes	302	Yes	Yes	Yes	301	1	58656.0	128717.0	4	23	Yes	Q2
59	No	306	Yes	Yes	Yes	390	3	55830.0	132373.0	4	22	Yes	Q2
29	Yes	332	Yes	Yes	Yes	340	3	25597.0	164431.0	6	33	No	Q2
47	Yes	406	Yes	Yes	No	342	5	93554.0	124100.0	4	49	Yes	Q2
41	No	424	No	Yes	Yes	314	2	72555.0	189339.0	5	22	No	Q2
62	Yes	434	Yes	Yes	No	332	4	77692.0	177025.0	3	10	Yes	Q2
49	No	445	No	Yes	Yes	368	1	61751.0	157607.0	5	14	No	Q2
60	Yes	458	No	No	No	330	2	38704.0	131236.0	5	17	No	Q2
35	Yes	464	Yes	No	No	318	5	33097.0	127955.0	3	33	Yes	Q2
25	No	466	Yes	No	Yes	370	3	76578.0	139846.0	5	15	Yes	Q2
42	No	499	No	Yes	No	352	3	49205.0	137812.0	6	22	No	Q2

Figure 7. Sample Report Generation

The reports generated can be represented graphically for analysis purposes. The reports are graphically represented for year 2015-2018 for the number of selected applications for approval of loans for North and South Branch as shown in Figure 8 and Figure 9 respectively. The visualization can be studied and evaluated by the higher management to help them in framing policies and making decision.

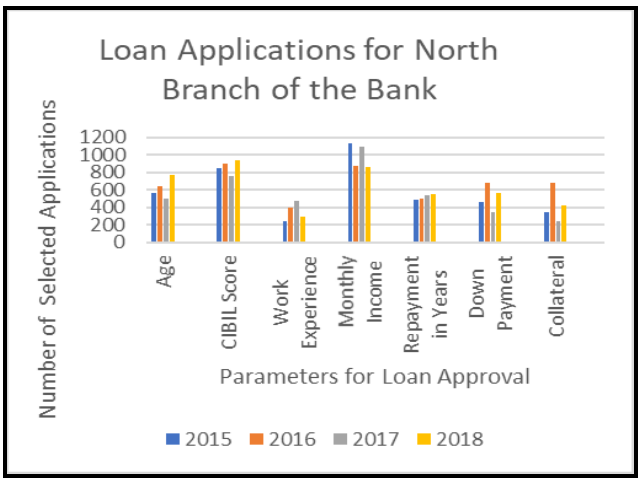


Figure 8. Result of Selected Loan Applications for North Branch of Bank

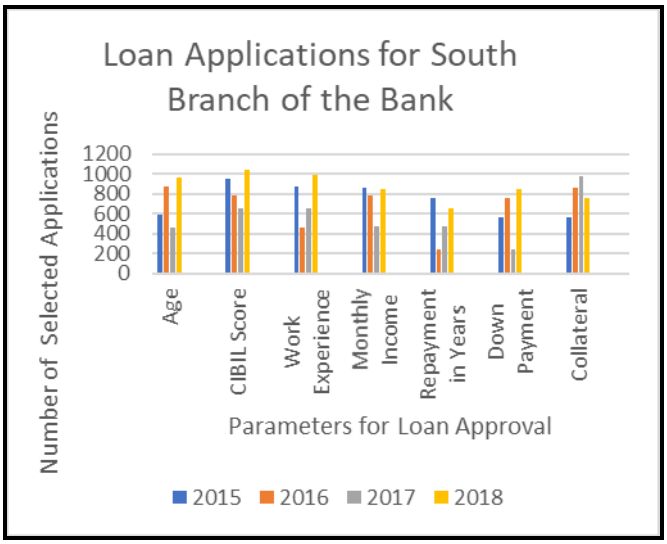


Figure 9. Result of Selected Loan Applications for North Branch of Bank

IX. RESULTS AND DISCUSSION

The data warehouse environment is more controlled and reliable for supporting business decisions. The data warehouse provides high quality information which is made available by cleaning the data effectively to remove inconsistencies in data. Only the valid data is transformed and summarized for computational purposes. Thus, it acts as single source of accurate and reliable information that can be utilized for data analysis. It is a collection of historical data basically used by the organization's management for decision making purposes. The huge volume of data when tested can result in generating knowledgeable outcomes which can be highly beneficial to the organization and shall help the higher management in making strategic decisions. Few outcomes can be predicted by analyzing the graphical representation as shown in Figure 8 and Figure 9. The outcomes are:

1. The number of customers interested in taking loans have increased, thus contributing in the increase of revenue and profitability of the bank.
2. For the span of four years, the customers are increasing who are maintaining their Credit Score.
3. The customers applying for loans at older age has increased over the years which depicts good economic condition of the country.
4. The North branch of the bank have customers with less work experience as compared to South branch of the bank.
5. As the North branch of the bank has customers of less working experience applying for loan thus the statics of monthly income is also less as compared to the South branch. This directly affects the down payment capacity also of the customers. Thus the down payment capacity of the North branch customers is lower than the South branch. Thus, it can be forecasted that the North branch of the bank consists of loan applications of customers of younger age group as compared to the South branch of the bank. Thus, the monthly income of the young customers is also less and proportionally the down payment capacity is also lower as compared to the customers of the South branch. Thus schemes to increase loan applications for the North branch can be undertaken to motivate and attract more customers in the North branch of the bank. Similarly, higher management can forecast and predict various other outcomes from the bank data warehouse. Analysts can make business decisions by analyzing the reports and pattern trends in the graphs. Management can modify existing policies and procedures to increase the growth of the bank and can have a healthy competition with their competitors.

X. CONCLUSION

Testing plays a vital role in a data warehouse. Data warehouse simply consolidates historical data at one place and is used to support strategic and business decision making. It helps any enterprise in making business decisions for the welfare and growth of the organization. Data warehouse works as a knowledge worker in the role of decision making and data analysis. It helps the managers to resolve problems within a defined timeline and facilitate effective decision making. The bank model developed can help the loan department of the banks to increase their customer potential and capability of applying for loan. It can also help the bank to identify their reserve requirements which is an important factor in granting loans to the customers. Data warehouses reports can help the banks in find out the performance of each branch and what are the opportunities to improve the performance. It can help the bank to identify more parameters to discover more knowledge from the bank data warehouse. Future work can focus on generating more test cases to execute on the data warehouse and results can be analysed to further enhance banking practices and operations. The focus of the future work will be analysis on prediction of Non Performing Assets (NPA) and Loan-To-Value (LTV) of loan department of the banks.

REFERENCES

1. 1. W. H Inmon, "Building the Data Warehouse", (Wiley, New York), 1996.
2. S. Nilakanta, K. Scheibe & A. Rai, "Dimensional issues in agricultural data warehouse designs", Computers and electronics in agriculture, 60(2), 2008, pp. 263-278.
3. B. Inmon, "DW 2.0; Architecture for the Next Generation of Data Warehousing", Information Management, 16(4), 2008, 8.
4. H. G. Hwang, C. Y. Ku, D. C. Yen, & C. C. Cheng, C. C, "Critical factors influencing the adoption of data warehouse technology: a study of the banking industry in Taiwan", Decision Support Systems, 37(1), 2004, pp. 1-21.
5. N. Balaban, & Ž. Ristić, "Business intelligence", Subotica: University of Economics Subotica, 2006.
6. B. A. Devlin, & P. T. Murphy, "An architecture for a business and information system", IBM systems Journal, 27(1), 1988, pp. 60-80.
7. C. Ballard, D. Herreman, D. Schau, R. Bell, E, Kim, & A. Valencic. " Data modeling techniques for data warehousing ", (pp. 25), 1998, IBM Corporation International Technical Support Organization.
8. S. Chaudhuri, & U. Dayal, " An overview of data warehousing and OLAP technology", ACM Sigmod record, 26(1), 1997, pp. 65-74.
9. M. Pathak, S. Singh, & S. S. Oberoi, "Impact of Data Warehousing and Data Mining in Decision Making", International Journal of Computer Science and Information Technologies, 4(6), 2013, pp. 995-999.
10. P. Ponniah, "Data warehousing fundamentals for IT professionals", John Wiley & Sons, 2011.

11. R. Khan, & S. Quadri, "Business Intelligence: An Integrated Approach", Business Intelligence Journal, 5(1), 2012, pp. 64-70.
12. R. A. Khan, & S. M. K. Quadri, "Business intelligence: an integrated approach", Business Intelligence Journal, 5(1), 2012, pp. 64-70.
13. E. Turban, R. Sharda, J. E. Aronson, & D. King, "Business intelligence: A managerial approach", (pp. 58-59), 2008, Upper Saddle River, NJ: Pearson Prentice Hall.
14. G. Kapfhammer, "The Computer Science Handbook", Chapter Software Testing, 2004
15. S. Mathur, V. Bali, S. L. Gupta, & P. Pahwa, " Data Warehouse Testing and Security: A Conspectus", International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(9S), 2019, pp. 843-850.
16. CTG, CTG Data Warehouse Testing, 2002
17. Mathen, M. P. (2010). Data warehouse testing. Infosys DeveloperIQ Magazine, 1-8.
18. Munshi, A. (2003). Testing a Data Warehouse Application. Wipro Technologies.
19. Krawatzeck, R., Tetzner, A., & Dinter, B. (2015). An Evaluation of Open Source Unit Testing Tools Suitable for Data Warehouse Testing. In PACIS (p. 22).
20. Tanuska, P., Moravcik, O., Vazan, P., & Miksa, F. (2009). The Proposal of Data Warehouse Testing Activities. In Central European Conference on Information and Intelligent Systems (p. 7). Faculty of Organization and Informatics Varazdin.
21. Golfarelli, M., & Rizzi, S. (2011). Data warehouse testing: A prototype-based methodology. Information and Software Technology, 53(11), 1183-1198.
22. ElGamal, N., ElBastawissy, A., & Galal-Edeen, G. (2013). Data warehouse testing. In Proceedings of the Joint EDBT/ICDT 2013 Workshops (pp. 1-8). ACM.
23. Kaur, P., Sharma, M., & Mittal, M. (2018). Big data and machine learning based secure healthcare framework. Procedia computer science, 132, 1049-1059.
24. Singh, R., Sharma, M., & Sharma, S. (2012). Sanjeev Kumar kaushal,"Management Information System".
25. Schutte, S., Ariyachandra, T., & Frolick, M. (2011). Test-Driven Development of Data Warehouses. International Journal of Business Intelligence Research (IJBIR), 2(1), 64-73.
26. Bateman, C. (2002). Where are the Articles on Data Warehouse Testing and Validation Strategy. Information Management.
27. Bhat, S. (2007). Data Warehouse Testing-Practical. Stick Minds.
28. Brahmkshatriya, K. (2007). Data warehouse testing. Stick Minds.

AUTHORS PROFILE

Ms. Sonali Mathur, working in JSS Academy of Technical Education, is a research scholar in Computer Science and Engineering Department at Birla Institute of Technology, Ranchi. She has received her M.Tech (Information Technology) degree from Guru Gobind Singh Indraprastha University. Her areas of interest are Data Warehousing, Data Mining and Database Management Systems.

Prof. (Dr.) S.L Gupta, working as Dean, Waljat College of Applied Sciences in Sultanate of Oman in Department of Management is an Educationist and Professor of Marketing Management. He has an excellent record to his credit right from the very beginning. His experience into academics has been very long. He has been teaching Marketing Management, Marketing Research, Sales and Distribution Management and Service Marketing since 1992. He believes in upliftment of deserving through micro finance.

Dr. Payal Pahwa, working as Principal and Professor, Computer Science & Engineering, BPIT, Rohini. She has more than 60 research publications in high quality National and International Journals/Conferences. She has membership of various professional societies.