# Prototype Data Warehouse to Analyze Foreign Currency transactions for Business Growth in Banks

***D.P.V.Kasinadh **Dr.P.Radha Krishna *** A.R.Joshi **** Arijit Laha**

## Abstract

*Data Warehouse helps the decision makers of an organization in taking decisions that helps in improving the business and profitability by consolidating data from many heterogeneous sources. In this paper, we worked on foreign exchange transactions, which is one of the important business lines of the bank in terms of profitability. We developed a model for analysis of these transactions using a multi-dimensional model especially star schema. ETL transformations have been designed based on the business requirements in order to convert operational data pertaining to foreign currency transactions into the proposed model. In this work we also build a cube and analyzed against the business requirements for which the model has been designed .*

## Introduction

After liberalization, in India foreign exchange is no longer a shortcoming, there is no short fall in supply. For banks there are new business potentials and opportunities in foreign exchange business; there is increased openness in Foreign Tourism, new services and foreign investments. At the same time there is an increased competition from other players like non-banking operators. In this situation banks really need to focus on the area in which there is a potential for growth and profitability.

A data warehouse helps management and decision makers turn raw data into valuable information and knowledge [6][12]. It gives management a systematic understanding of events, which lets them redefine or reengineer an organization's business processes for competitive advantage [4] [8]. It can be used for data analysis in different kinds of applications like retail, financial services, healthcare etc. The most commonly mentioned benefit of data warehousing is simplicity as it provides a single view of the data.

In banks everyday large volume of data is produced as a result of activities, and as a by-product of various transactions. This information about the customers and their activities can be used for business growth. Yet, most of these data is not used properly to generate information necessary to support strategic decision-making [12]. Once the data warehouse is implemented and customized to business and IT needs, helps bank in understanding the need of its customers and become more competitive.

** Associate Professor, IDRBT, Castle Hills, Road No. 1,Masab Tank, Hyderabad-500057, Andhra predesh, E-Mail: prkrishna@idrbt.ac.in.

However, data warehouses are expensive, require extensive planning and are resource intensive by nature. Building a data warehouse is a challenging issue, because compared to software engineering it is relatively new discipline and does not yet offer well-established strategies and techniques for the development process. A lot of projects in this area fail due to the complexity of the development process [11].

In this work we designed a star schema for analyzing the foreign currency transactions and created a cube to do analysis. The remainder of this paper is structured as follows: a brief discussion of related implementations is presented in section 2. Section 3 presents a brief overview of Foreign Exchange services and data warehousing. Requirement analysis and data model is described in section 4. Section 5 describes results and we conclude in section 6.

## Related work

Building a data warehouse is a well-studied problem. For this purpose various methods have been proposed and widely used. In this section, we reviewed and presented a brief discussion of some of the works in building data warehouse for business development in different fields.

Cunnigham *et al* [5] outlined the process of building data warehouse for CRM; it explains the need for maintaining relationship with existing customers rather than trying to acquire the new customers. Next it presents the need for techniques, technologies and concepts of knowledge Management, Data Mining and Data Warehousing in order to support the organization's relationships with its customers, and presents a robust multidimensional *starter* model that supports CRM analysis. This model is used to analyze various profitability analyses such as customer profitability analysis and market profitability analysis.

Du and Wong [3] addressed the design of data warehouses for supply chain partners. It describes importance of data warehouse in the supply chain environment as it a complex system, which involves the passing of information across the companies that present different levels of layers of the system. This study proposes five approaches to designing data warehouses for supply chain partners, namely Centralized Data Warehouse, Distributed Data Warehouse, Coordinated Data Warehouse and Federated Data Warehouse and Heterogeneous Data Warehouse and Presents advantage and shortcomings of each and every approach.

Lawyer and Chowdary [10] presents the data warehousing architecture and practices used at a major U. S. retailing company. The paper discusses the two predominant styles in data warehousing, namely the "Bill Inmon Style"(or the top-down approach) and the "Ralph Kimball Style" (or the bottom-up approach) and merits and demerits of each approach. It finally explains the steps that are needed to be followed in data modeling, keys, Loading and challenges encountered during the process of creation of warehouse related to data and technology.

In [13], Singhal and Jajodia presented how data warehouse and data mining techniques can improve the performance and usability of Intrusion Detection System in order to identify usage patterns and catch fraudulent activities, as these techniques can automatically analyze the data. It also presents a technique to model network traffic data using a star schema.

Benefits of employing data warehousing techniques and On Line Analytical Processing (OLAP) tools in preparing data for estimating the input parameters of various decision models was presented by Vasilakis, El-darzi and Chountas[14]. They proposed a relational and a multidimensional data model for storing simulation output data and describes analytical requirements and implementation issues of the proposed model. Here a snowflake schema has been used rather than a star schema because of performance related issues.

In this paper our focus is on use of data warehousing techniques to help a Bank in improving its performance and profitability. We want to design a multidimensional model and to the analysis on one of the important business lines of the bank i.e., foreign exchange services.

## Background

In this section we presented description about foreign exchange services provided by banks and the importance of data warehousing in this area.

## Foreign exchange service

Foreign exchange, in simple terms means foreign currency. In India, foreign exchange has been given a statutory definition, it states, " all deposits, Credits and balances payable in any foreign currency and any drafts, traveler's cheques, letters of credit and bills of exchange, expressed or drawn in Indian currency, but payable in any foreign currency" [2].

Foreign exchange services is very important as lot of currency is traded every day Worldwide, can influence the lives of billions of people [1]. Foreign Exchange exists as a result of i) Trade and investment ii) Speculation and iii) Hedging.

Companies, which import or export, goods or services are buying them in one currency and selling them in another currency, it means they pay out money in one currency and receive money in another. So companies need to convert some of the money they receive into the currency in which they pay for goods. Similarly a company that buys an asset in a foreign country has to pay for it in the local currency of the foreign country and so will need to convert its home currency into the local foreign currency or into the currency the counter party wants. In this process banks help their customers in acquiring the required foreign currency. Thus banks buy and sell the currencies that their customers are lending or receiving.

In India, banks are most active players in foreign exchange Trade. They have a wide network of branches and corresponding banks all over the world. Generally commercial banks act as intermediary between exporter and importers who are situated in different countries. These banks undertake foreign exchange transactions by getting license from Reserve Bank of India (RBI). Banks who are given license to deal in foreign exchange are called Authorized Dealers.

To maximize the profits and to be competitive, knowledge-workers have to take quick and efficient decisions. To achieve this easy and quick access to aggregated and consolidated information is required.

## Data Warehouse

Data Warehouse is built by integrating data from multiple and heterogeneous data sources and it contains summarized and aggregated data on various pre-determined dimensions. To facilitate complex analysis and visualization, the data in a warehouse is typically modeled as multidimensional, consists of two parts a) A large central table called fact table containing the bulk of data and b) A set of smaller dimension tables one for each dimension.

The multidimensional schema resembles a star, with the dimension tables displayed in a radial pattern around the central fact table. Granularity of the fact table is important while designing as it shows the lowest level of information that is stored in the fact table. Granularity mainly depends on two factors i) Dimensions to include and ii) Hierarchy of the dimensions. It is up to the user to decide the granularity he/she requires, i.e., on the kind of analysis to be performed [13]. *Dimensions* are the different perspectives for an entity that an organization is interested in. A dimension table contains the analytic criteria by which the data warehouse is queried. *Facts* are numerical measures and they can be thought of as quantities by which we want to analyze relationships between dimensions. The *fact table* contains the names of the facts as well as keys to each of the related dimension tables. [9]. Fig.1 shows the architecture of a data warehouse. It includes various tools for extracting data from multiple sources of data, for cleaning, transforming and integration of this data. Data in warehouse is present in one or more servers, which present a multidimensional view of data to a variety of front-end applications like query/reporting tools, analysis applications and data mining applications. Data Warehouse provides the users with an online analytical processing (OLAP) facility by which they can query the data warehouse. To overcome the cost and the time involved in aggregating large amount of data for a frequently accessed query, we store data in a multidimensional cube.

These cubes provide a very quick response to the users.

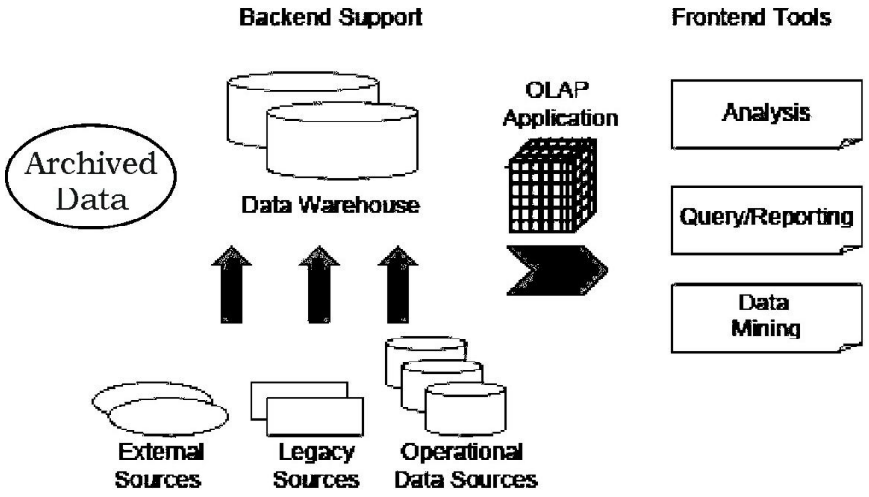


**Fig.1 Data Warehouse Architecture ([4])**

If the foreign currency transactions are properly organized into a multidimensional star schema, it is possible to do analysis. From this analysis, for a bank it is possible to understand the patterns present in the transactions in terms of what currency is having high demand at what time of the year; which branch's performance was good over the last six months? So that higher authorities can provide special privileges to that branch. And with which country a correspondent branch can be established. And bank can quickly know the surplus amount currency wise and take appropriate decisions.

## Modeling Foreign Exchange transactions

RBI collects data pertaining to Foreign Exchange Transactions from Banks at regular time intervals. This is done through the Electronic Reporting System; format of reporting is common to all Banks. This information is used by RBI to assess the balance of payments (BOP), it is important as it affects the position of foreign currency. This information gives a fair idea of various dimensions of foreign exchange business that are useful to design the dimensional model. Structure of the data is available at RBI website [15]. Data related to foreign currency transactions is present in the form of master tables and transaction tables. We used master tables for creating the dimensions of interest and a transaction table to create the fact table.

In this paper, our focus is on foreign exchange currency service provided by the Bank. The model we designed here is based on the transactions that are held at different branches of the bank.

First step in designing the schema for the given problem is to identify what are different kinds of analyses that are relevant to foreign exchange services?. We identified two categories of analyses namely (a) Business performance (b) Management information systems (MIS), which are relevant to foreign currency transactions and these validated by business expert. Here we are giving some of these questions in each category.

(a) Business Performance: These questions need to be answered by the dimensional model in order to find its performance / profitability over the different branches, regions, and time periods and in different currencies by performing analysis on the data present in the data warehouse.

(i). To identify the performance (business) of branches over time.

(ii). What are the currencies with high volumes of transactions and low volume of transactions? That is ranking of the currencies based on the volume of transactions of respective kind.

(iii). Analysis based on purpose. For each purpose what is the volume of the business.

(iv). What is time and region wise volume of transactions in different currencies? At what part of the year a particular currency is in large volume of transactions and in which region.

(b) MIS: Following are for providing information to the higher authorities.

(i). Average sales / purchase of each currency (region, Time).

(ii). Which branches have no sale / purchase in the last month?

We developed a model to organize the raw data (transactions) into Star Schema so that it is easy to query and analyze. The Transactions can be viewed as a multidimensional data. Granularity of the fact table used here is individual transactions, i.e. fact table contains information about each interaction with the customer. For the analysis of foreign exchange transactions we identified five dimensions namely Geography, Time, Purpose, Country and Currency.

Required dimensions are created by using ETL (Extract, Load and Transform) operations from Master data. The ETL operations performed on the input data include 1) Extraction of data from input sources. 2) Merging of data from different tables. 3) Validation of the fields. 4) Checking for empty of Null or invalid values and filtering of those records or replacing with default values. 5) Generation of Surrogates Keys for dimensions. 6) Joining of the fact and dimension tables. We modeled these ETL operations using SAS ETL studio.

Geography dimension describes the location of the branch where the transaction took place. This dimension consists of a hierarchy, which shows how this information can be aggregated to produce different views (Region wise, City wise, Branch wise). Time dimension

contains the date, fortnight, month, quarter, and year of the transaction; it provides different views of timing like daily, monthly, and quarterly. Purpose dimension describes the type of the transaction whether it is sale or purchase of the Currency and reason (actual purpose) of the transaction. Currency dimension describes the type of currency of the transaction. Country gives the description of the country from which the money is coming (Purchase) or going (Sale). Fig.2 shows the hierarchy present in all the five dimensions.

The attribute in the fact table "Amount in Foreign Currency" is a measure, which is amount of foreign Currency (Described by currency dimension). Another measure is "Amount in Indian Rupees" is the actual amount in Indian rupees for which the transaction was held. Fig 3 shows the Star schema. A popular abstraction for multidimensional data that is widely used in OLAP is the data cube. A cube is simply a multidimensional structure that contains at each point an aggregate value, i.e. the result of applying an aggregate function to an underlying relation.

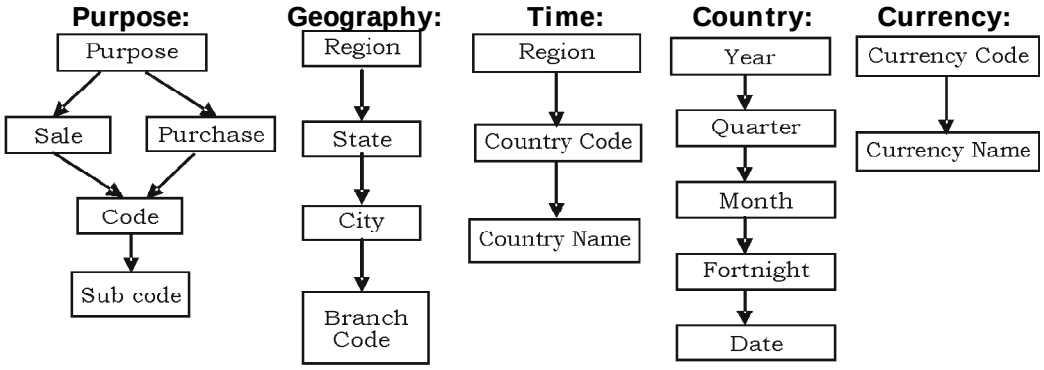


**Fig.2 Dimensions and Hierarchy**

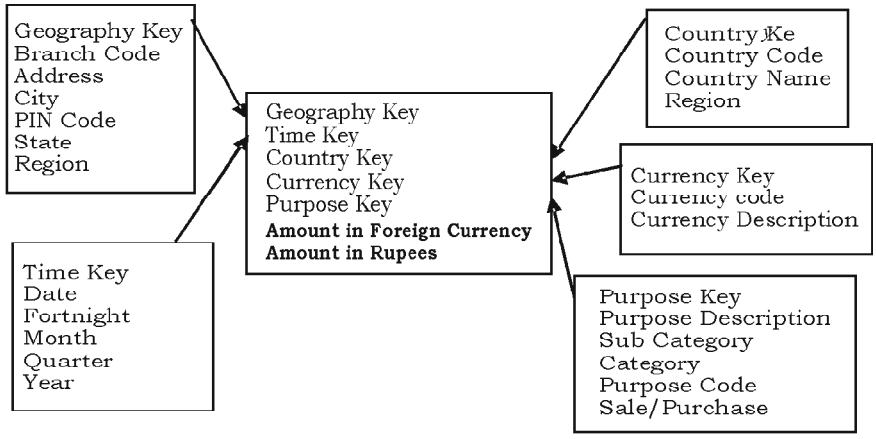


**Fig.3 Star Scheme for Forex Transactions**

Using the schema shown in Fig 3, a corresponding cube would be a five dimensional structure, in which a cell contains measures of the facts. For example a cell may correspond to a transaction on a particular date in particular currency, it can be sale or purchase to a particular country and the transaction held at a specific branch of the bank.

## Analysis and discussion

The proposed data warehouse is implemented in MS-SQL Server as database running on Windows XP, the hardware was IBM NetVista PC 2.4 GHz processor; the schema was populated with sample fact table records and five dimensions Queries for analysis are written in multidimensional expressions (MDX), a language specifically designed for OLAP analysis. Each query could be modified to include additional measure by adding fields from fact table and the dimensions. One of the queries for requirements is discussed below.

The following MDX statement can be used to figure out the performance of a branch, over a given time period in terms of measure "amount in Indian rupees".

SELECT NON EMPTY Hierarchize({ [Time].[AllTime].[2003],[Time].[All Time].[2003].Children,[Time].[All Time].[2003].[qtr3].Children, [Time].[All Time].[2003].[qtr4].Children } ) ON COLUMNS, NONEMPTY {[ Geo]. [AllGeo]. [HYD].[ANDHRA].[BHIMAVARAM].[TBH4007] } ON ROWS FROM [Forex_New] WHERE ([Measures].[Amount Rs] )

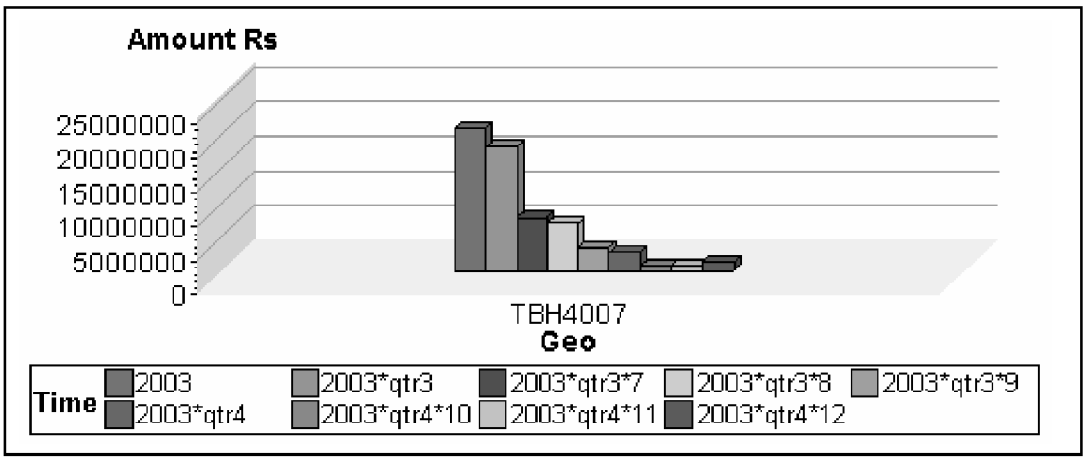


Fig. 4 Visualization of results of query

Fig.4 shows the output of the MDX query given above in the form of a graph. Amount of business in rupees is represented on Y-axis and Time is represented on X-axis. It shows the total business (sales and purchases) of the branch TBH4007 over third and fourth quarter of the year 2003 over all purposes and currencies. From the graph it is possible to observe the trends present in the amount of the business.

## Conclusion

In this paper, we have focused on building the data cube for Foreign exchange currency transactions of a bank. Here we designed a prototype multidimensional model based on the master data and transactional data. This data model is used to perform the analyze business performance, MIS.

Our on-going work is to provide decision maker with qualitative analysis by using multi-attribute summarization and data fuzzification.

## References


- An introduction to foreign exchange and Money Markets. The Reuters financial training series.1999.
- H.P.Bhardwaj. Foreign Exchange Handbook. Bhardwaj Publishing Co. 1998
- Timon C. Du, Jacqueline Wong. Designing Data Warehouses for Supply Chain Management. International Conference on E-Commerce Technology, IEEE 2004. Pp.1-9
- S.Chaudhuri , U.Dayal. An overview of data warehousing and OLAP technology. SIGMOD Record, 1997.
- Colleen Cunnighan, II-Yeol Song, Peter P.Chen. Data Warehouse design to support customer relationship management analyses. DOLAP-2004, pp. 14-22
- Gerri Furlow, The case for building a Data Warehouse, IT Pro, IEEE, 2001 pp.31-34.
- Stella Gatziu, Manfred Jeusfeldy, Martin Staudtz , Yannis Vassilioux .Design and Management of Data Warehouses. Report on the DMDW'-99 Workshop
- Matteo Golfarelli, Stefano Rizzi, Iuris Cella , Beyond Data Warehousing: What's Next in Business Intelligence? . DOLAP-2004, pp.1-6.
- Ralph Kimball, The Data Warehouse Tool kit. Wiley Publications (1996).
- Jeff Lawyer, Shamsul Chowdhury. Best Practices in Data Warehousing to Support Business Initiatives and Needs. Pro. in the 37th Hawaii International Conference on System Sciences – IEEE 2004
- Beate Listl, Robert M. Bruckner, Karl Machaczek, Josef Schiefer, A Comparison of Data Warehouse Development Methodologies Case Study of the Process Warehouse, DEXA 2002, LNCS 2453, pp. 203–215, Springer-Verlag Berlin Heidelberg 2002
- P. Radha Krishna. An Approach to Establish Data Warehouse for Banks in India, The ICFAI Journal of Systems Management 2004. pp. 17-43.

- Anoop Singhal, Sushil Jajodia. Data warehousing and data mining techniques for intrusion detection systems. Distributed Parallel Databases Springer science+Business Media, LLC 2006.pp. 149-166
- Christos Vasilakis, Elia El-Darzi, Panagiotis Chountas. A Data Warehouse Environment for Storing and Analyzing Simulation Output data. Proceedings of the 2004 Winter Simulation Conference IEEE 2004.
- http://www.rbi.org.in/Scripts/ElectronicReportingSystem.aspx.