

Developing Analytical Data Mart For Financial Markets

Fathima Begum F.R. ^{#1}, M.V.sarvesha ^{#2}, Y.S.KumaraSwamy ^{#3},

^{#1}Lecturer, King Khalid University, Abha, Kingdom of Saudi Arabia, fathimasafar7862003@yahoo.com

^{#2}ResearchScholar ,CMJ University,Meghalaya,INDIA,sarvesh_mv@yahoo.com

^{#3} Department Of CSE,NCET,Bangaluru,yskld2swamy@yahoo.com

Abstract: In this paper we describe the evolution and architecture of data mart developed to address the modelling and analysis needs of financial markets. More specifically the data mart is used in projects relating to price trend analysis and historical characteristics of the companies. The primary focus of the data mart is on the detailed movement of prices/volumes in the last ten years and different financial characteristics of last five years, has organised in the different periods. We show that the data mart goes well beyond consolidating data from different sources, like yahoo, Nse India and MSF by including a number of complex pre calculated fields like different moving averages to the needs of technical analysis. A prototype system for performance analysis has been developed and the results are discussed.

I. INTRODUCTION

A datamart is a relational database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data, but it can include data from other sources. It separates analysis workload from transaction workload and enables an organization to consolidate data from several sources.

In addition to a relational database, a data mart environment includes an extraction, transportation, transformation, and loading (ETL) solution, an online analytical processing (OLAP) engine, client analysis tools, and other applications that manage the process of gathering data and delivering it to business users. A common way of defining data mart or warehousing is to refer to the characteristics of a data warehouse are subject oriented, Integrated, Nonvolatile and time Variant as set forth by Inmon [1]:

In the financial markets around the world, trading in the stock market has gained extraordinary popularity as a way of life to reap huge profits. However, this generates the large amount of data which poses problems in standardization of data, accessing the information and computation of indexes to academicians and practitioners.

Given the pressure of the modern state of business perhaps, there is, nothing more valuable than time. In view of this companies developing data marts to analyze large volumes of data generated every day in the financial markets. To identify the different sectors and stocks in focus by these data there are several factors that have driven this change. The primary factor is to understand the individual stocks performances and forecasting the prices based on the history. The end goal of this is to minimize the time to access the data on stocks and their previous performances.

In order to track all of this information we have begun to develop analytical data marts which are useful in financial markets. Analytical data marts differ from traditional data marts in the sense that they focus on capturing, monitoring and storing key performance indicator's (KPI'S) that links the analytics. Instead of storing all of the details behind the

calculations, analytical data marts and warehouses simply store the results of the calculations. The another difference is that traditional data marts focus on explaining the characteristics of stock while analytical data marts focus on tracking and monitoring the stock performance on the surface it may appear that this is a small distinction, but it has profound implications on the data mart

II. DATAMART ARCHITECTURE

The architecture of a datamart by necessity is complex, and includes many elements. The reason for this is that a datamart is an amalgamation of many different systems. Integration of diverse elements is its primary concern, and to accomplish this integration, many different systems and processes are necessary.

The datamart consists of the following architectural components, which compose the datamart infrastructure [3]:

- *Technical architecture:* Hardware, software, network, database management system.
- *Data models:* Dimensional.
- *Metadata layer:* Data about data. This includes, but is not limited to, definitions and descriptions of data items and business rules.
- *Data acquisition:* The process of loading data from the various sources.
- *Data distribution:* The dissemination/replication of data to applications for specific segmented groups.
- *User analysis:* Includes the infrastructure required to support user queries and analysis.

Figure 1 shows the overall view of a data warehouse with each of these components.

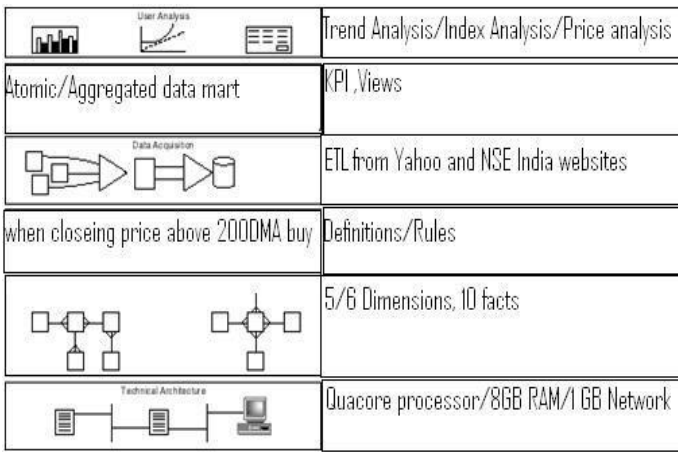


Fig 1. Datamart architecture

III. EXPERIMENTS

(i) Data Description

The data used in this paper was obtained from the different websites like yahoo finance,nseindia and MSN. The daily data of the each stock with the attributes open, low ,high and close prices and fundamental data like ROC, ROA EPS, P/E ratio etc of 40 quarters are downloaded from the open sources. The datamart consists of the daily stock prices of the 700 stocks of 10 years from the year jan'2000 to the aug'2011 and also fundamental data of each stock of 40 quarters .

(ii) Preparing the Datamart

Preparing a data mart involves more than reading columns from a source table to the data mart. It also includes theaggregation or transposition of observations from "multiple-row-per-subject" tables like transactional tables and timehistories. This process is a critical success factor for being able to answer the business question or to have goodpredictors available for a target event or target value.The missing values are filled by using the average of preceding and succeeding values. In this paper the data mart is created as "One row percustomer"[4].

The datamart start with the combining of the data from different data tables , the daily stock prices of the stock are included between the period 01/01/2000 to 26/08/2011 and fundamental data is included in quarterly .The different moving averages like 3DMA,5DMA,10DMA, to 200DMA of the each stock is calculated stored in the datamart. In creating the datamart the lot of variables in the data are transformed in to required format and irrelevant variables are deleted and target variables DMAs are calculated.

The dimensional model is of the data mart is shown in fig2 below . The data mart six different tables which consists of different information about the individual stocks based on the sector and industry type ,the fact_ta table consists of the daily data and fact_fa table consists of fundamental data , Time table data consists of characteristics time data

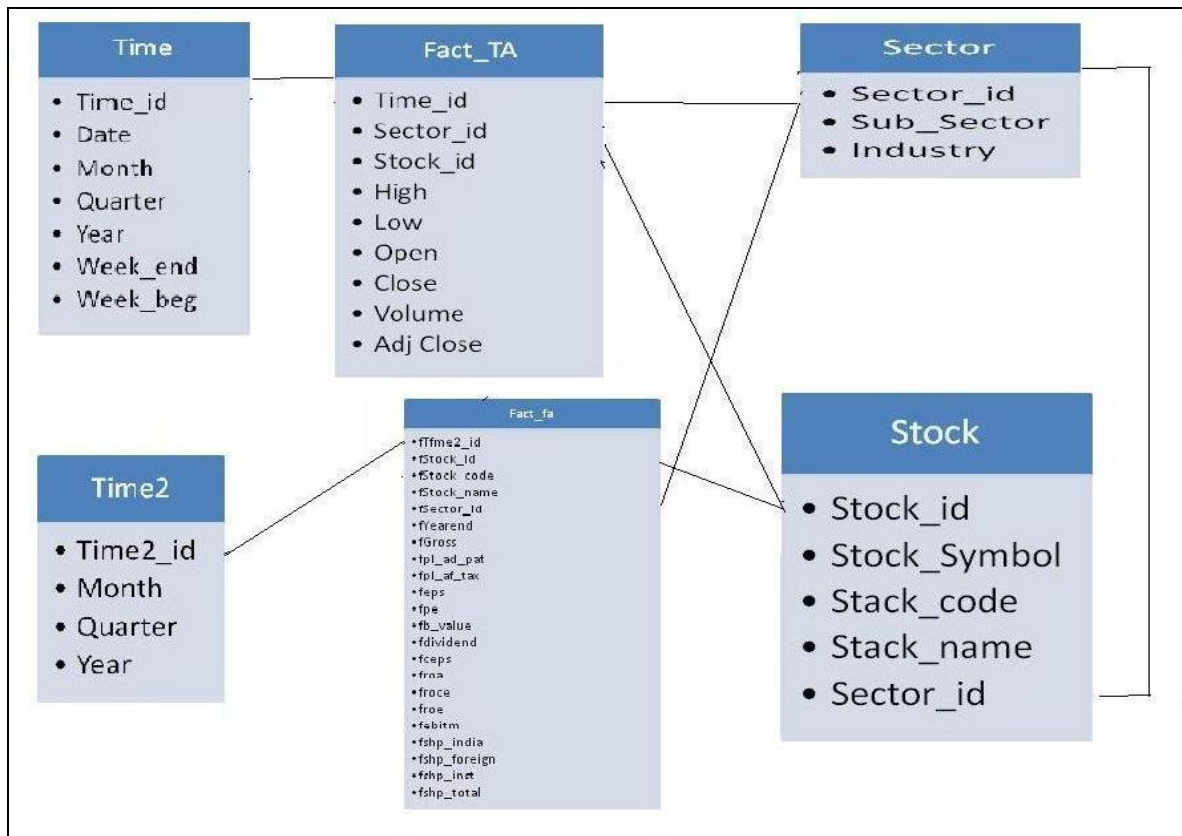


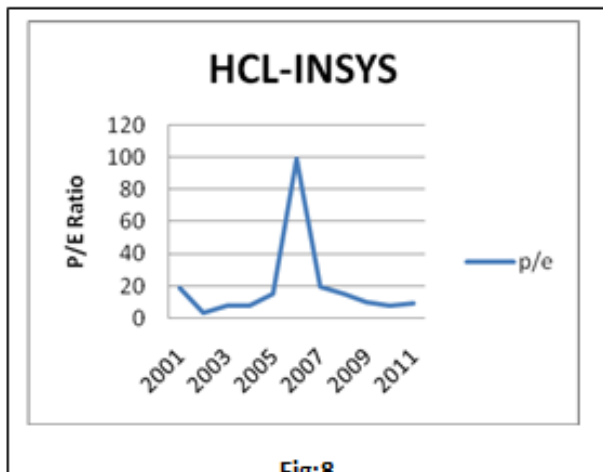
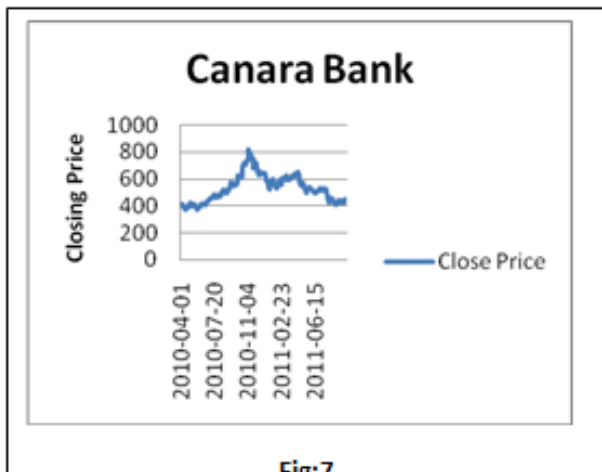
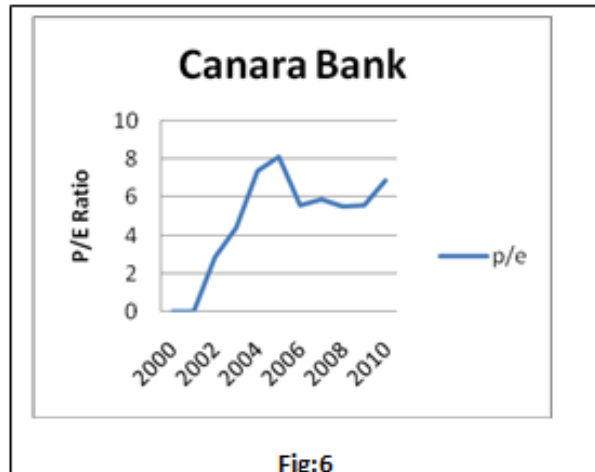
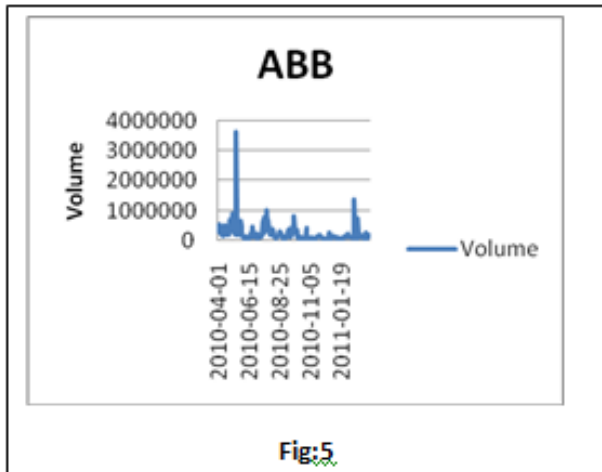
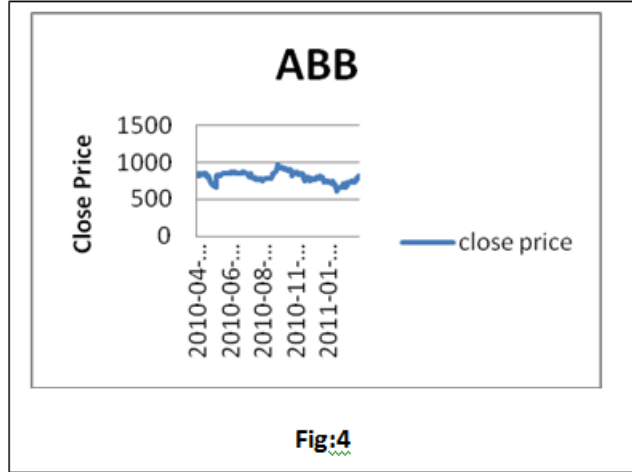
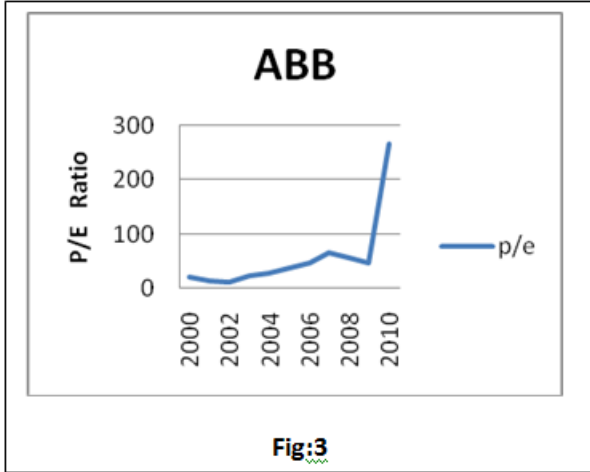
Fig2: Design Of DataMart

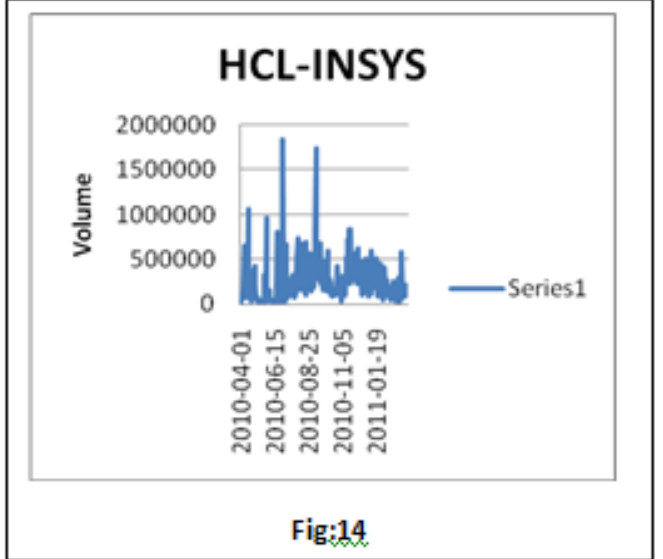
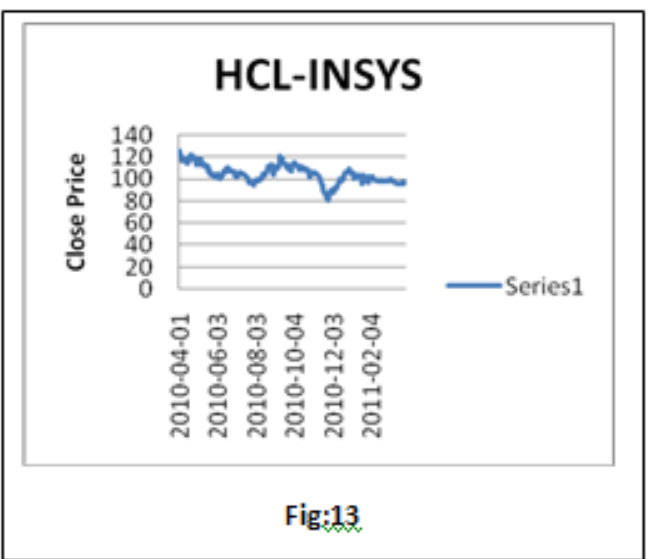
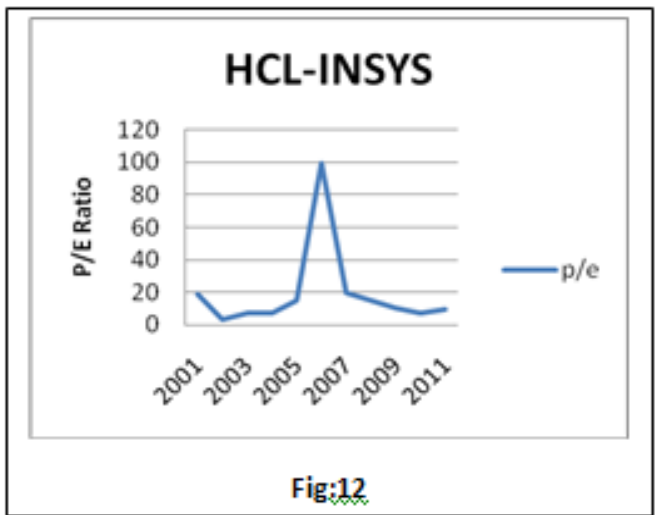
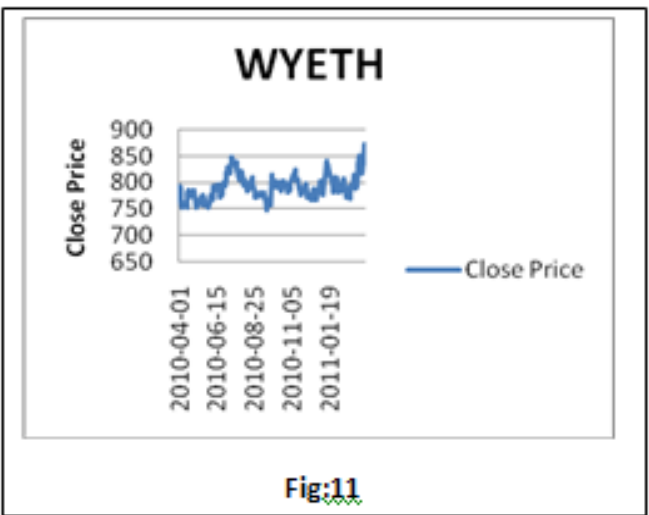
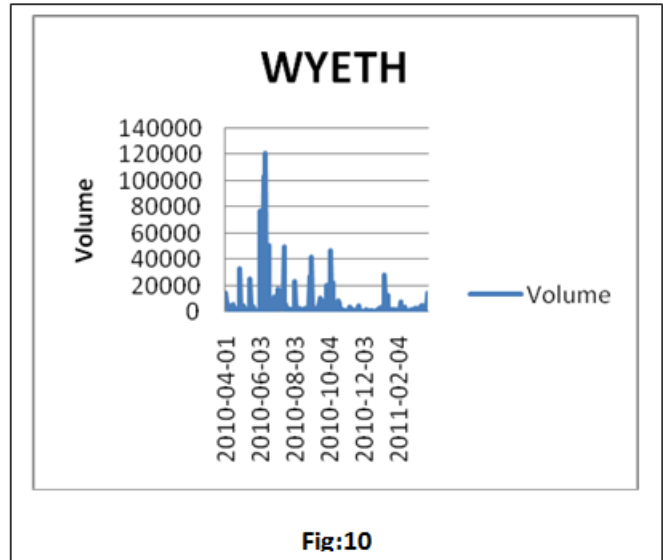
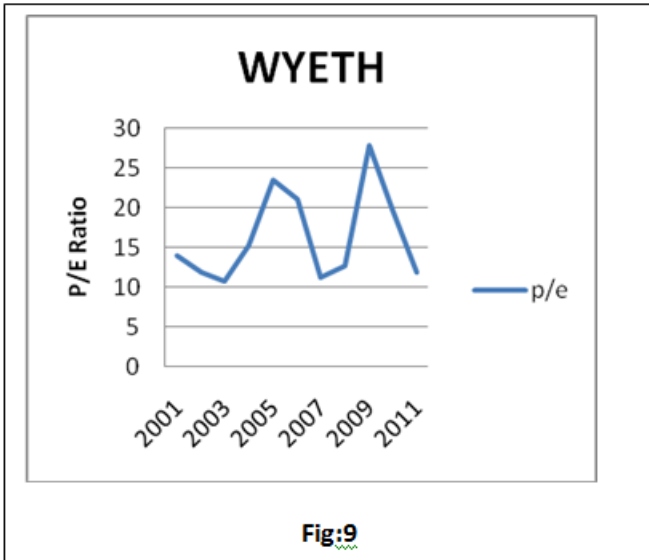
The research study is aimed to propose a system model for customer centric eGovernment system. Then, a suitable modeling formal language will be proposed to represent system. The applicability of model will be illustrated using typical system structures. The analysis techniques and properties of modeling language will be used to analyse

dynamic properties of system. The proposed model can be used for visualization to non-technical users, system analysts and developers to understand the system clearly. The design and development of information system with this model can be constructed with selected technology platform.

IV. RESULTS

The graphical trend analysis for the few stocks are given in the following graphs which are taken from the data mart







IV. CONCLUSIONS

The financial markets requires the systems that gives the nature and strength of the stock based on the analysis, this leads to the creation of the data mart which supports the technical and fundamental analysis of the stock .In this we have created the data mart which supports both fundamental and technical analysis of about 500 stocks. Further this can be extended to test the different rules which leads to identify the profit giving stocks

V. REFERENCES

1. Inmon C Mastering Data Warehouse Design : Relational and Dimensional Techniques, John Wiley,2003.
2. Jay B.Simha (2005) Developing a Decision Support System for Telco using Dimensional Datamart.
3. O'Neil B, Oracle data warehousing unleashed, Sams publishing,2000.
4. Gerhard Svolba, "Efficient One-Row-per-Subject Data Mart Construction for Data Mining", paper 078-31, SUGI31.
5. Inmon B, "The problem with Dimensional Modeling", DM Review, May 2000.
6. Reeve L.L "What is a Business Dimensional Model?", DM Review, Nov 2003.
7. Mattison, R Data Warehousing Strategies, Technologies and Techniques. New York:McGrawHill,1996.
8. James A Cox ,"Mining the Data mart: A case Study with Stock Market Data", SAS Institute Inc, NC
9. Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications*, 31, 515–524.
10. Wei, C. P., & Chiu, I. T. (2002). Turning telecommunications call details to churn prediction: A data mining approach. *Expert Systems with Applications*, 23, 103–112.
11. Ralph Kimball (2008) The Data Warehouse Toolkit.