

**Delivering on the BI Promise:
A Financial Services Roadmap**

Nishith Singh
DecisionStudio Software Services Pvt Ltd
April 2006

Abstract

The Financial Services Industry in India has been relatively slow in the adoption of data-mining and Business Intelligence (BI). To meet the Basel II requirements and to do business better in today's hyper-competitive environment Financial Services need to jump start their analytics efforts immediately.

This paper emphasizes the need to set up the right data infrastructure for BI initiatives through setting up of data warehouse and departmental data marts. As a road map towards doing comprehensive analytics, this paper lists down and prioritizes various projects that are doable right now and can deliver demonstrable business value very early in the implementation cycle.

By using open source software for deploying BI, organizations can reduce costs, improve ROI, and reduce the overall implementation risk for the BI initiative.

About the Author

Nishith Singh is a Director at **DecisionStudio** Software Services Pvt. Ltd. (<http://decisionstudio.com>), a Bangalore based analytics consulting and services firm that helps its clients deliver on the BI promise through innovative use of open source software. Nishith has prior analytics experience with i2 Technologies, CitiFinancial and GE Consumer Finance Analytics, and can be reached at nsingh@decisionstudio.com. His personal blog on doing advanced analytics using open source software is at <http://opensourceanalytics.com>, and his articles can be found on **IT Toolbox** portal at <http://blogs.ittoolbox.com/bi/opensource>.

Table of Contents

ABSTRACT	1
ABOUT THE AUTHOR	1
TABLE OF CONTENTS	2
INTRODUCTION	3
ANALYTICS AND BASEL II	3
THE CASE FOR BI, DATA-MINING, ANALYTICS	4
DATA WAREHOUSE – THE FOUNDATION FOR SUCCESSFUL BI	6
A ROADMAP FOR FINANCIAL SERVICES	7
<i>Marketing Data Mart</i>	7
<i>Adding Customer and Product Information</i>	8
<i>Sales and Credit</i>	8
<i>Portfolio Data Mart</i>	8
<i>Financial Data</i>	9
<i>Balanced Scorecard & Business Performance Management (BPM)</i>	10
PLANNING FOR A SUCCESSFUL IMPLEMENTATION	10
THE CASE FOR OPEN SOURCE	11
ENSURING A SUCCESSFUL IMPLEMENTATION	11
CONCLUSION	12

Introduction

Globally the Financial Services industry has been an early adopter of data warehousing, data-mining, analytics, and Business Intelligence (BI). Large financial institutions have known that competitive advantage in business is dependent on gathering data, analyzing it, and generating detailed information on products, customers, channels and staff. With the advent of tighter regulatory controls and the need to meet Basel II requirements, analytics and BI have graduated from being just a competitive advantage to an imperative.

However, the use of data-mining and analytics by the Indian Financial Services firms has so far been relatively limited. Some have dabbled in analytics through expensive one-shot efforts that did not grow into full-blown implementations, and a few others believe that they can do without analytics - counting solely on the long list of standard and ad-hoc reports provided by their IT departments. Prohibitively expensive BI software coupled with the shortage of trained resources further complicates the situation.

Yet the need for doing analytics and BI can no longer be ignored. Analytics and BI, when done properly, can create immense benefits for the business along multiple dimensions.

This paper lays down a road map for rolling out analytics in a phased iterative manner without incurring million dollar expenses. The projects outlined in this paper are doable today. We go deeper into the details of how they can be done and have also outlined some best practices that would help in seamless execution.

Analytics and Basel II

In June 1999, the Basel Committee issued a proposal for a New Capital Adequacy which has to be implemented by 2006. This is known as the second Basel Accord or, more commonly, as Basel II. Basel II provides for a framework based on three "mutually reinforcing pillars". These are:

- 1. Minimum capital requirements.** Banks need to either record, measure and control credit and operational risks, or else increase their capital reserves.
- 2. Supervisory review process.** Each financial institution needs to adopt effective internal processes for comprehensive evaluation of its risks to assess capital adequacy. Regulatory supervisors need to ensure this and intervene as necessary.

simpler terms, data-mining is what you do when you are unable to *know* your customers as you would if you were living and working in a small community.

It is also a lot about sifting through the piles of data to identify and flag events and patterns of note, and making them available to the people at the right time. BI is what can allow the top management to get a clear view through the fog of war.

While the terms data-mining, BI, and analytics are used rather interchangeably, they all are essentially about extracting implicit, previously unknown, and potentially useful information from large data sets. The Analytics process can thus be considered the conversion of raw data into optimal and actionable decisions.

Activities and Roles:

- **Data Warehousing and Dimensional Modeling:** Standardization of data for easy retrieval. Provides a single version of truth, and reduces “cost-per-analysis”. *A well-designed Data Warehouse is the foundation for a successful BI/Analytics initiative.*
- **Reporting, Cube Analysis, & What-If Analyses:** Analysis of historical data to understand drivers and to measure, control, and manage better
- **Data Mining and Predictive Modeling:** Using patterns hidden in historical data to predict future behavior
- **Information Delivery for Decision Support:** Making sure the right people get the information needed through reports, scorecards, dashboards and Executive Information Systems (EIS).
- **Automation of Operational Decisions:** Bridging IT with analytics to automate operational decisions (such as flagging potential frauds for verification, or sending pre-defined communication to potential churns)

Business Rules and Strategies can be defined by mining and modeling of customer data. These strategies can be used for automating operational decisions where predictive information comes in handy for overall process optimization. For example, a fraud detection system can keep routing potential frauds to a verification queue. The fraud scoring model ensures that only the applicants with highest likelihood of fraud get queued to the verification group.

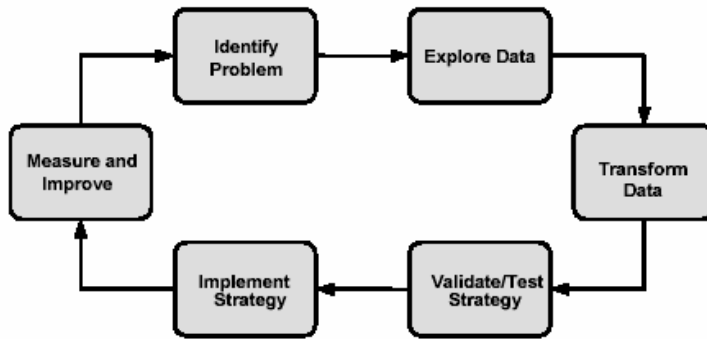


Fig. 2: Deployment of Strategies

Models have a lifetime and after a while their efficiency starts to drop and they need to be either retrained or completely replaced by a newer and better strategy. **Champion-Challenger** evaluation of competing strategies should be used for managing the model and strategy lifecycle.

Data Warehouse – The Foundation for Successful BI

The Data Warehouse is the basic foundation all successful BI initiatives are built upon. If an organization is looking at long term sustainable analytics as a competitive advantage, building a data warehouse is an absolute necessity that cannot be skipped. Some direct benefits of keeping data in a Data Warehouse include:

- Contains Preprocessed, Directly Usable Data
- Single Repository for Multiple Data Sources
- Easily Available - SQL
- Avoids “multiple versions” of “truth”
- Reduces Cost-Per-Analysis
- Reduces Cycle Time for Idea to Execution

The primary difference between an application database and a data warehouse is that while the former is designed (and optimized) to record, the latter has to be designed (and optimized) to respond to analysis questions that are critical for business.

Application databases are **OLTP** (On-Line Transaction Processing) systems where every transaction has to be recorded quickly, and these systems are **write-optimized**. A Data Warehouse on the other hand is a database that is designed for facilitating querying and analysis. Often designed as **OLAP** (On-Line Analytical

Processing) systems, these databases contain read-only data that can be queried and analyzed efficiently. In this sense an OLAP system is designed to be **read-optimized**.

Data Mart versus Data Warehouse

A **Data Mart** is a specific, subject oriented, repository of data designed to answer specific questions for a specific set of users. So an organization could have multiple data marts serving the needs of marketing, sales, operations, collections, etc. A data mart usually is organized as one dimensional model as a star-schema (OLAP cube) made of a fact table and multiple dimension tables.

In contrast, a **Data Warehouse (DW)** is a *single organizational repository* of enterprise wide data across many or all subject areas. The Data Warehouse is the authoritative repository of all the fact and dimension data (that is also available in the data marts) at an atomic level.

From an implementation perspective, it is better to build individual data marts with an eye upon the Data Warehouse as recommended by Ralph Kimball – a Titan of Data Warehousing. This allows you to start delivering upon business needs and demonstrating the value of the BI initiative in a quick iteration of limited complexity.

A Roadmap for Financial Services

Setting up of a Data Warehouse for a Financial Services company need not be a cause of heart-burn and anxiety for the IT department. By following an iterative phased approach of developing data marts, while keeping long term objective of setting up an enterprise data warehouse, it is possible to deliver on the BI promise.

The enterprise data warehouse can be flexibly built over time in a phased manner to provide the infrastructure for multiple reporting and analysis needs including: Financials, Profitability, Relationship Management and Marketing, Credit, Risk, Regulatory Compliance, Asset Liability Management, Process Optimizations, and Performance Management.

Marketing Data Mart

If you are starting from scratch, marketing could be a good place to begin. You can construct a fairly simple data mart containing marketing reach-outs (promotions, campaigns, mailers) and response information. With this information in place, marketing can now do Promo and Campaign Analysis, compute response rates, and run test-marketing campaigns for better targeting of marketing effort.

Sample Data: PromoID, CustomerID, DateTime, ResponseFlag, Product, Geography, Medium, etc.

Adding Customer and Product Information

A straight-forward step after deploying the Marketing Data Mart is to pull in detailed information about customers and product/service offering, making sure that it can be linked to the Marketing Data Mart through unique identifiers (CustomerID and ProductID). This gives your marketing analysts the power to link up customer's marketing behavior back to Customer Demographics and Product/Service Characteristics. They can now do Response Modeling for optimization of marketing efforts, and target specific customer segments with specific fine-tuned marketing campaigns.

Sample Data: Customer Demographics (Age, Location, Sex, Education, Marital Status etc.); Product Details such as Financial parameters, Type, Category and Sub-category.

Sales and Credit

Now that you have data flowing in about your marketing efforts, the logical next step would be to focus on your sales conversion process. This would include information on each product/service sale, sale parameters, and credit information. With this in place you can start generating Sales and Credit Reports directly from the data warehouse, and also generate reports on cross-sell and up-sell, new acquisitions, credit approvals etc by customer and market segments.

Sample Data: ApplicationID, CustomerID, ApplicationDate, AuthorizationDate, Credit Decision, CreditLimit, DisbursalDate, Disbursal Amt, InterestRate/APR, IRR, Charges, Tenure, EMI, FraudFlag, etc.

It would be prudent to take a breather at this stage and review the results of your BI efforts at this point. There may be reconciliation issues that need to be investigated and fixed at source. This is also a good time to take all round user feedback and measure the ROI on the project. It is extremely important to drive the adoption of the Data Warehouse and build user confidence on the numbers coming from it.

Portfolio Data Mart

Armed with a few successful initiatives, and having solved any issues in the previous stages, now is the time to move on to bigger things and start looking at financials.

The first step towards building a complete Financial Data Mart is to pull in at least the monthly account-level portfolio data. Account level historical data allows you to generate a wide variety of portfolio and regulatory reports, measure delinquency and default at account and ad-hoc segment levels, and monitor trends and roll-rates. It also gives your analysts direct access to account level behavioral, financial and credit data, thus laying the foundation for a solid risk management practice. This information can then be used to do risk-based pricing, predict potential delinquents, and adopt optimal methods for curing delinquent accounts.

At this stage you also have sufficient information to start credit scoring and fraud prediction to proactively approach credit risk, and to compute exposure and Value at Risk at a very fine granularity as required for capital adequacy under Basel II.

Sample Data:

1. **Loans:** RecordDate, LoanID, Product, Loan Amt, AssetCost, AuthDate, Tenure, Int Rate, EMI, CustomerID, GuarantorID, Co-borrowerID, Principal Outstanding, Interest Outstanding, Delinquent POS, Bucket, Charges, InterestIncome, DelinquencyString, etc.
2. **Bank Accounts:** RecordDate, AcctID, Product, Amt, StartDate, Month on Books, OverdraftFlag, Overdraft Interest Rate, CustomerID, Charges, InterestIncome, etc.
3. **Credit Cards:** RecordDate, AcctID, CardNo, IssueDate, ValidityDate, APR, Product, CurrentOS, LastBill Amt, LastBillDate, Last Paid Amt, Last Paid Amt, Delinquent OS, Bucket, ChargesIncome, InterestIncome, DelinquencyString, etc.

It would be prudent at this point to make sure that the portfolio/account financials in the Data Warehouse reconcile back to the relevant financials in the ledger. If there are discrepancies, they need to be investigated and reconciled. Taking care of data integrity and reconciliation issues at each step goes a long way in generating user confidence on the numbers coming out of the Data Warehouse.

Financial Data

The logical next step is to pull in the rest of financial information to complete the overall financial picture of the enterprise. This is easier said than done with this information often scattered across multiple systems and departments. It can be tough to identify these sources and then to rope them in but the payoff is immense and immediate.

At this stage you have to basically aim at bringing in all the revenues and expenses at an account level and then try to map them back to individual portfolio accounts. This may require Unit Costing and Transfer Pricing of Activities to ensure that expenses are allocated back to individual loan/account/credit card as accurately as possible. It is possible to use assumed values but then the assumptions should be recorded and periodically validated.

It is now possible for analysts to generate Product Profitability, Yield Analysis and Expense Analysis at portfolio level as well as for ad-hoc segments. Coupled with dimensional analysis (OLAP cubes), this can be an extremely potent tool for identifying and catering to profitable and under-served market segments. Financial and Profitability analysis capabilities for ad-hoc segments make it possible to do modeling and forecasting of Default Risk and Performance at Portfolio, Segment and Account Level. Armed with the amount of information now available, organizations can gradually try out Business Experimentation which was made famous by Capital One.

Balanced Scorecard & Business Performance Management (BPM)

BPM can be used to efficiently track and manage performance of processes like Customer Service, Payment/Cheque Processing, Collections, Repossessions, etc. Each of these may require identification and pulling in of new types of data from multiple sources. Once data is available for a particular process, it can then be used for performance management by defining and evaluating KPIs.

A Balanced Scorecard summarizes carefully chosen key metrics into a single report that aids execution and implementation of business strategy.

Planning for a Successful Implementation

With the business benefits and expenses at stake in a BI rollout, it is critical to plan in advance and decide upon the overall execution strategy. This involves prioritization of projects, identification of business sponsors and functional team members, gaining internal buy-in, and getting the required budget approvals.

Equally critical is the vendor and technology selection process and to make sure that the vendor understands your domain and business. It is not uncommon to find vendors who may know a lot about the BI software but close to nothing about your business. Ideally you should conduct a proper vendor evaluation and choose from multiple options.

Some key things to be kept in mind during a BI vendor evaluation:

- **Selection Criteria:** Decide upon measurable and demonstrable criteria for the evaluation..
- **Drive the Evaluation:** Make sure that the vendor is not driving the evaluation process for you or distracting continuously.
- **Check Reference Sites:** Not all the “references” may be happy/satisfied customers. Some actually are extremely unhappy, just that they cannot say it publicly.
- **Understand Financial & Licensing Issues:** Yes, commercial OLAP total cost of ownership (TCO) goes into multi-million dollars.
- **Proof of Concept:** It is extremely important to compare short listed vendors on a small project.

The Case for Open Source

While choosing the BI software it is important to understand that the first year software license fee is going to be only a small part of the overall project cost. People often overlook recurring annual license costs, costs of vendor and technology lock-in, and resource costs which can dramatically impact the TCO for the project.

Open Source Software provides an excellent option for doing comprehensive BI by providing extensive capabilities, greater flexibility, easier customization, and much higher ROI than the commercial software. Open source also makes an ideal candidate for iterative development under constantly changing requirements as there are no sunk licensing costs. We very strongly recommend using open source software in order to control costs and manage the risk inherent in such projects.

There are companies that provide end-to-end services including consulting, solution design, construction, deployment, and ongoing support. DecisionStudio is one such consulting firm focused exclusively on delivering BI solutions and services based on fine-tuned open source software.

Ensuring a Successful Implementation

The BI implementation is likely to be the most iterative project undertaken by your group. Being a key component of long term business strategy, the analytics team needs to be flexible and should work in close coordination with the functional teams.

Iterative, bottom-up development is required and the traditional waterfall development model could be counter-productive. It is important to start small, deploy quickly, and iterate frequently, and it may be a good idea to apply some of

the techniques of **Agile Development** for BI implementation projects. The team needs to be flexible and would have to relentlessly manage scope. Any change to a leg of the **Requirements-Resource-Time Triangle** (due to misunderstood requirement, less time or fewer people) requires a corresponding change to at least one other leg.

Some critical steps to keep in mind for a successful BI implementation:

- **Identify Business sponsors** and gain their support.
- **Define** your business requirements, critical success factors, data sources and business rules.
- **Design** the data marts in the context of a wider enterprise data warehouse: Develop the overall framework, strategy and technology for an enterprise data warehouse built bottom-up with incremental data marts.
- Use **open source software** and agile **iterative development**: It can be extremely risky to purchase multi-million dollar commercial software in a project where requirements will change. Open source provides the same capabilities at dramatically lower ownership cost.
- **Use prototypes, early and often**: Prototypes make it easier for business users to visualize the proposed solution.
- Use **external resources** and **business partners**: Engage support from software, hardware, and consulting partners for a successful BI implementation.

Conclusion

The Financial Services industry faces new challenges resulting from intense competition across all sectors and more complex regulatory requirements. Competitive advantage in today's business is heavily dependent upon gathering detailed internal data and acquiring external data. Sophisticated analyses against the data can help increase customer profitability, improve service offering, manage risk, and reduce operating costs.

It is possible to do comprehensive analytics, data-mining and BI today using open source software to reduce project risks and increase the ROI. With a focus on users' needs and by using open source software in a phased deployment, the IT departments can deliver upon the BI promise.