Chapter XIII
Data Warehousing and Analytics in Banking: Implementation

L. Venkat Narayanan
Satyam Computer Services, Singapore

ABSTRACT

Data warehousing and analytics represent two of the foremost technologies that can be used by banks to obtain sustainable competitive advantage. Adopting the right implementation methodology is critical to ensuring successful implementation. Alternate implementation methodologies, typical challenges in implementation, and critical success factors apart from real-life case studies are discussed here as learning points to aid in successful implementations. Future proofing implementations are critical to avoid rework, and hence some key emerging trends have also been discussed.

INTRODUCTION

Business users in banking use data warehouse and analytics as key enablers to achieve sustainable competitive advantage. While there is considerable debate on the definition of success as well as on the various surveys that provide statistics on the failure rate, there is no argument on the fact that data warehouse and analytics implementations have the potential to fail if inappropriately implemented. The author advocates the view that success of a data warehouse and analytics implementation should be measured by the positive impact to the top line and bottom lines of the bank instead of assessing technical factors alone. Hence a technically successful data warehouse that has the right data, made available at the right time to the right users is still a failure if it does not end up increasing the top line and/or bottom line.

CHALLENGES IN IMPLEMENTATION

Data warehousing is often touted as a high-risk implementation. According to many surveys, a large majority of data warehouse implementa-
Data Warehousing and Analytics in Banking: Implementation

Implementations are considered failures. The definition of failure and the authenticity of the survey results can be a point of debate. However, it is generally accepted that data warehouse implementations can go wrong more often than not. Some of the key challenges in data warehouse implementations are listed below, with specific reference to banking implementations:

- **Ability to define user requirements**: Defining the requirements for a data warehouse becomes a challenge when considering the maturity levels of the users. This is because, unlike an operational system, the users do not have an 'old system' or manual equivalent to refer back to, to help them define what would be best suited to their current requirements. This being a relatively new paradigm, not all users are mature enough to define what they want. One potential mitigation option is to use packaged data warehouses that would encapsulate the best practices in terms of banking data models and reporting templates. Vendors like IBM, Teradata, Oracle, and Satyam offer such packaged data warehouse solutions.

- **Data reconciliation**: Since the data warehouse is also used for regulatory and compliance initiatives, there is a need for validating the accuracy of loaded data. This is different from data warehousing when applied in other industries where accuracy can be conceded to an extent, as most analysis happens at aggregate levels. The recommended method of data reconciliation and validation for banking data warehouses is to validate the loaded data against general ledger balances. For example, the outstanding balance amount of loan accounts should be matched against the general ledger chart of the account into which the loan account balances are posted. This requires the ability to identify the charts of accounts into which key measures are posted in the general ledger.

- **Poor data quality**: Data quality is a major challenge in banks, as the data elements that are valuable in a data warehouse environment like customer demography are not considered necessary for day-to-day banking operations. From the operations perspective, what are needed are customers’ identification and contact details, which will allow the bank to contact the customers and send statements. For instance, data items like annual income are considered unnecessary from the operations perspective, whereas they are very critical in the data warehouse environment to develop predictive models. With the push towards customer centricity in banks, there is a greater thrust to get clean data in banks. This challenge is mitigated by a combination of business rules to clean data when it is brought to the data warehouse (English, 1999) and change in business processes to get more updated and clean data.

- **Lack of organizational readiness**: Too many banks failed to get the business benefit of data warehousing, upon treating it as an IT project. Business users must be mature enough to use the data warehouse to achieve their business objectives and goals. This requires a change in culture and business processes. If the change management in culture, process, and compensation is not addressed as part of implementation, it would lead to a failed implementation, with ROI hardly visible (Adelman, Bischoff, & Dyché, 2002).

- **Long timeline**: In many banks, IT departments treat data warehouse implementations as infrastructure initiatives resulting in long project timelines. The resultant impact is that the users lose interest and develop alternative options. The right way to implement the data warehouse is to break into granular implementation tracks, with each track tied to a specific business objective or need. The ideal timeline for each iteration