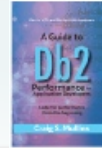




Mullins Consulting, Inc.
The Web Site of Craig S. Mullins



A Guide to Db2 Application Performance for Developers
By Craig S. Mullins
Order now!
A new book to help programmers write efficient Db2 code
Covers both Db2 for z/OS and LUW

[Home](#)

[Services](#)

[Articles](#)

[Presentations](#)

[Books](#)

[Speaking 2021](#)

[Social Media](#)

[Database Links](#)

[Contact Us](#)

TERADATA.
MAGAZINE ONLINE



The Complete Guide to DBA Practices and Procedures
By Craig S. Mullins
Second Edition Available Soon!

Harness The Full Power of Your Data

by Craig S. Mullins

December 2006

A large wireless phone service provider was concerned with the number of customers it was losing. Every customer lost costs the company \$53 in monthly revenue. Although the revenue looks small on a customer by customer basis, with a large customer base the company was losing millions of dollars each month. Using advanced analytics they were able to develop an attrition model to predict which customers were most likely to terminate their contract. In doing so, the company developed a model to cross-sell helping them to retain customers by providing products, services and other incentives targeted to their profile. This program improved the retention rate and contributed to an overall savings of \$6.7 million.

That is the type of success story common among companies that have deployed advanced analytics to better understand their data. Advanced analytics is a business-focused approach, comprising techniques that help build models and simulations to create scenarios, understand realities, and future states. Advanced analytics utilizes data mining, predictive analytics, applied analytics, statistics and other approaches in order to follow organizations to improve their business performance.

Traditional business intelligence enables us to understand the here and now, and even some of the why, of a given business situation. Advanced analytics goes deeper into the "why" of the situation, and delivers likely outcomes. Although advanced analytics cannot infallibly predict the future, it can provide models for judging the likelihood of events. By allowing business managers to be aware of likely outcomes, advanced analytics can help to improve business decision-making with an understanding of the effect those decisions may have in the near future.

Advanced analytical capabilities can be used to drive a wide range of applications, from operational applications such as fraud detection to strategic analysis such as customer segmentation. Regardless of the applications, advanced analytics provides intelligence in the form of predictions, descriptions, scores, and profiles that help businesses better understand customer behavior and business trends.

Issues in Deploying Advanced Analytics

When implementing advanced analytics projects it is not uncommon to encounter problems along the way. One of the potential difficulties involves managing and utilizing large volumes of data. Businesses today are gathering and storing more data than ever before. New data is created during customer transactions and to support product development, marketing, and inventory. And many times additional data is purchased to augment existing business data. This explosion in the amount of data being stored is one of the driving forces behind analytics. The more data that can be processed and analyzed, the better the advanced analysis can be at finding useful patterns and predicting future behavior.

However, as data complexity and volumes grow, so does the cost of building analytic models. Before real modeling can happen, organizations with large data volumes face the major challenge of getting their data into a form from which they can extract real business information. One of the most time-consuming steps of analytic development is preparing the data. In many cases, data is extracted, and a subset of this data is used to create the analytic data set where these subsets are joined together, merged, aggregated, and transformed. In general, more data is better for advanced analytics. There are two aspects to "more data": (1) data can increase in depth (more customers, transactions, etc.), and (2) data can grow in width (where subject areas are added to enhance the analytic model). At any rate, as the amount of data expands, the analytical modeling process can elongate. Clearly performance can be an issue.

Real-time analytics is another interesting issue to consider. The adjective real-time refers to a level of responsiveness that is immediate or nearly immediate. Market forces, customer requirements, governmental regulations, and technology changes collectively conspire to ensure that data that is not up-to-date is not acceptable. As a result, today's leading organizations are constantly working to improve operations and with access to and analysis of real-time data.

For example, consider a financial services provider that is confronted with detecting and preventing fraud. Each transaction must be analyzed to determine its validity. The retailer waits for approval while this is done in real-time. But if you err on the side of safety, valid transactions may be declined which will cut into profit and perhaps more importantly, upset your customer. The advanced analytics approach leverages predictive analysis to scrutinize current transactions along with historical data to ensure transactions that may appear suspicious aren't the norm for this customer. The challenge is doing this in real-time.

Today's nimble organizations need to assess and respond to events in real-time based on up-to-date and accurate information, rules, and analyses. Real-time analytics is the use of, or the capacity to use, all available enterprise data and resources when they are needed. If, at the moment information is created (or soon thereafter) in operational systems, it is sensed and acted upon by an analytical process, real-time analytics have transpired.

As good as real-time analytics sounds, it is not without its challenges to implement. One such challenge is reducing the latency between data creation and when it is recognized by analytics processes.

Time-to-market issues can be another potential pitfall of an advanced analytics project. A large part of any analytical process is the work involved with gathering, cleansing, and manipulating data required as input to the final model or analysis. As much of 60% to 80% of the man-effort during a project goes toward these steps. This up-front work is essential though to the overall success of any advanced analytics project.

Let's move on and examine the solutions Teradata offers to these issues involved in implementing advanced analytics solutions.

Advanced Analytics Solutions

Teradata Warehouse Miner helps reduce the cost of analytics by pushing the data intensive tasks directly in the Teradata database. Teradata Warehouse Miner provides scalable enterprise analytic modeling technology that reduces data mining cycle time and provides faster delivery of information. Because it performs all of the analytic functions within the database, it significantly improves performance and provides the scalability required to build analytic models. Teradata Warehouse Miner also provides an interface that enables other tools to run models directly in the Teradata database and features to facilitate creation of the analytic data set required for all other advanced analytic modeling tools.

Teradata Warehouse Miner is a set of advanced analytic products that automate and optimize your enterprise analytic process that:

- Reduces cost and improves performance of the analytic model and model development by integrating analytics into your Teradata warehouse;
- Shortens time to results by automatically deploying models and associated data;
- Improves data quality through automated assessment and identification of questionable data, and;
- Empowers users to easily score customer or product segments using sophisticated models.

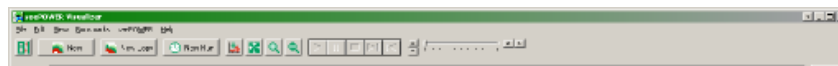
Furthermore, deploying best practices can enable a Teradata warehouse to be analyzed using statistical analysis tools such as SAS and SPSS to mine larger volumes of data. For example, a global entertainment company was running a SAS model to forecast the sales of new movie titles. Each forecast required a little over 7 minutes per title. This was reasonable until running the model against 300 titles translated to 36 hours. By splitting the task and pushing 90% into Teradata, the company was able to dramatically improve performance by a factor of 28.

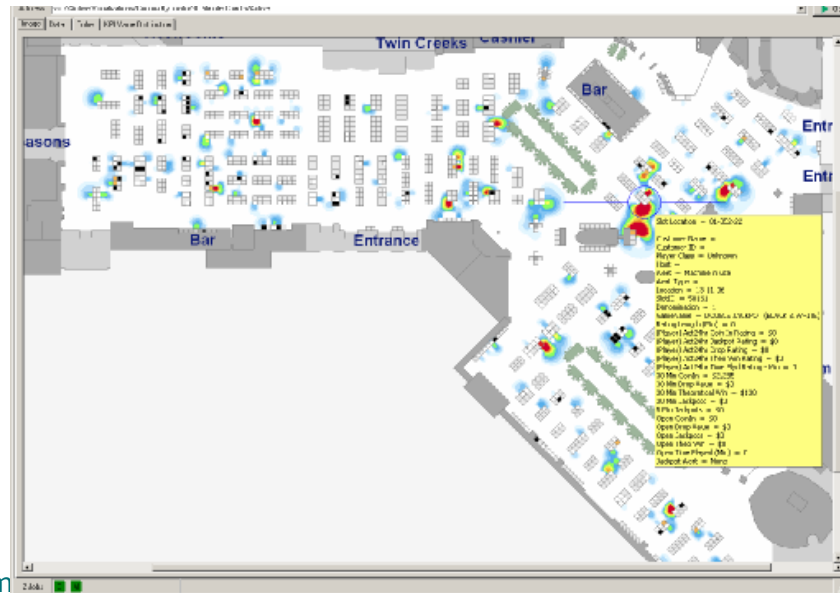
Additionally, Teradata supports real-time analytics projects with advanced visualization techniques. Compudigm, a Teradata business partner, provides the capability to transform massive volumes of transaction and customer data into critical, real-time visual insights to make smarter, faster and more profitable business decisions.

The combination of the Teradata warehouse and Compudigm's seePower software delivers the ability to query against atomic level data across the enterprise in real time delivering intuitive visualizations that can be viewed over space and time. By combining thousands of data points into a visual representation, business users can more rapidly glean trends and patterns from massive amounts of data – making information more readily available, digestible, and actionable.

By using seePower to analyze customer history in a Teradata warehouse you can score customer history and manage real-time activity. To analyze customer history in real-time, customer rankings (scoring) need to be determined to assure the proper level of service. These rankings are generally done in batch and used in real-time. Compudigm leverages this information to present the real-time picture of the business.

Teradata and Compudigm offer a particularly useful implementation of visualization and real-time analytics in the Gaming business. The key for real-time analytics in Gaming is delivering customer value. It is imperative that the focus of real-time analytics for Gaming capitalizes upon events and opportunities as they unfold on the floor. The name of the game is to be able to monitor and service top VIP customers as they play, where they play (refer to figure).





Visualization and real-time analytics enable casinos to have a rapid response to monitor and service VIPs, up-sell to sub-VIP customers while they are on the floor, and monitor hot players to proactively acquire new loyalty card members. Real-time advanced analytics coupled with visual and dynamical presentation results in better-served customers and a more profitable Gaming business.

And remember the speed to market issue? Teradata can help reduce time required to deliver on the many tasks required to support advanced analytics projects. Instead of requiring days or weeks for data extraction, joins, subsetting, merging, aggregation, and transformation, all these functions can be done directly in the Teradata Database against large data volumes. Teradata Warehouse Miner facilitates the setup of this analytic data set so SQL novices can quickly create an analytic data set optimized for Teradata. One customer built a ten million customer record with more than 500 different variables. On a server, this task took more than six hours; however, by moving the processing into Teradata Database, the process took 15 minutes.

Best practices are also important. When models supporting analytics projects are run infrequently, or there are only a few models, it makes sense to do this work as a part of each specific project. However, when an organization begins incorporating dozens or hundreds of models into their business environment on an ongoing basis, the repeated manipulation of large amounts of data becomes inefficient. Teradata provides best practices that greatly condense overall processing cycles and vastly reduce the time to create, update, or implement any given model.

As many models are built over time, certain standard metrics and manipulations will surely become apparent. For customer analysis, it is hard to imagine that total customer spending or number of customer transactions would not be of interest in most analysis efforts. Similarly, it is hard to imagine that total store sales for recent periods would not be of interest to most store level analytics. At the same time, any required cleansing or recoding of the detail data required to facilitate such roll ups will be constant once the right analytic procedure is established.

An Enterprise Analytic Data Set (ADS) takes the standard data rollups that are used in a variety of analytic tasks and centralizes their generation. Instead of each analyst or process having to incorporate all of the logic and consume all of the processing time needed to derive the data for each analysis, the standard metrics will be created in an automated fashion on a regular schedule and made available to all analysts and processes. Any entity that will be the focus of a wide range of analytics is a candidate for an Enterprise ADS.

This methodology improves consistency in the methodology that various analysts and processes use to generate their analytical data sets. And, the chance of an error being made due to the omission or altering of appropriate logic is removed. Overall system processing cycles are greatly reduced since variables requiring a lot of heavy processing will be computed once, stored, and shared, rather than being run time and again. Analysts can get straight to adding value with their work rather than focusing repeatedly on the same basic prep work. Basically, the Enterprise ADS provides a simplified view of the data warehouse by providing a condensed manageable number of analytic tables to view that represent hundreds of tables of detailed data.

A Customer Example

In 1994 Continental was ranked at the bottom in most airline customer satisfaction categories. Today, Continental has turned that around winning accolades for its business performance.

Continental didn't have a clear idea of who their customers were. The company had multiple definitions of "customer" and "value," but no data to define each. They were running dozens of databases with duplicate information and too often their highest-value customers were not identified and received lower compensation.

Continental deployed the Teradata Warehouse to build a compensation history and standardized compensation matrix, resulting in improved customer satisfaction and loyalty. It also identified where the customers fit on the value chain. Continental was able to identify three primary segments which aligned with its customer's profiles. The high-value customer who is willing to pay higher fare for personalized service, the elite customer who travels frequently and reacts to compensation, and the regular, infrequent customer.

Now, based on these profiles, if there's a flight delay of 90 minutes or greater, customers in the highest two segments are treated differently. Since compensation doesn't motivate the high-value customers, they receive a personalized letter apologizing for their inconvenience. On the other hand, Elite customers are motivated by compensation and are at higher risk of defecting if inconvenienced. They receive a complementary pass to the Presidents club. Other customers receive nothing.

The result is happier customers being treated individually by the airline and cost savings because compensation is only given to the customers who best react to it. Advanced analytics programs such as this have helped Continental to move from Worst to First to Favorite among US airlines.

The Bottom Line

The end result of deploying advanced analytics is increased productivity with the ability to gather and analyze large volumes of data to deliver faster, more-effective business decisions.

From [Teradata Magazine](#), December 2006.

© 2012 Craig S. Mullins,

DB2PORTAL.com

© 2021 Mullins Consulting, Inc. All Rights Reserved [Privacy Policy](#) [Contact Us](#)