



## SAS® Analytics Accelerator for Teradata

In-database analytics improve data governance on predictive analytics projects and produces faster, better results

### What does SAS® Analytics Accelerator for Teradata do?

SAS Analytics Accelerator for Teradata allows a core set of SAS statistical and analytical functions used for data summarization, discovery and predictive modeling to be executed directly within the database environment. It includes predictive modeling algorithms as well as the preliminary tasks of optional sampling, statistical transformations, variable selection and dimension reduction to derive better models.

### Why is SAS® Analytics Accelerator for Teradata important?

It will help organizations promote better data governance required for predictive analytics and achieve better analytic outcomes in an efficient manner. SAS Analytics Accelerator for Teradata supports a complete SAS model development process in Teradata without extracting the data from the database.

### For whom is SAS® Analytics Accelerator for Teradata designed?

SAS Analytics Accelerator for Teradata is specifically for SAS® Enterprise Miner™ users or users who develop programs leveraging SAS statistical procedures, such as those in SAS/STAT® or SAS/ETS® software. It is designed for analytic professionals (data miners, statisticians and analysts) to build predictive and descriptive models on the latest set of data residing in the database, enabling them to leverage their SAS skills and produce SAS results more quickly.

There are many reasons that in-database processing is making its way to the forefront of decision makers' minds. IT directors recognize that the cost of integrating and managing data across the enterprise is growing. Their staff members are spending increasing amounts of time preparing and transferring data, often duplicating efforts, rather than using an integrated environment to build effective predictive analytical models that could provide deeper insights for the organization. In addition, scattered data marts are maintenance headaches, and often analytic professionals are restricted to using small sets of data or data that is not up-to-date.

For their part, business leaders become frustrated when they cannot get timely answers, and they are recognizing that lengthy, inefficient analytical model development is affecting their organization, whether it is the inability to detect new incidences of fraud or gain and keep the most profitable customers.

SAS Analytics Accelerator for Teradata enables the execution of key SAS analytical tasks within a Teradata database or data warehouse. This type of in-database processing reduces the time needed to build, execute and deploy powerful predictive models. It also increases the utilization of the enterprise data warehouse or relational database to reduce costs and improve data governance that is required for successful analytic applications.

### Key Benefits

- **Reduce data movement and redundancy to ensure data quality and enhance resource use.** In-database analytics reduce, or elimi-

nate, the need to move massive data sets between a data warehouse and the SAS environment or other analytical data marts for multipass data preparation and compute-intensive analytic functions. Deploying data preparation tasks and analytics within the same data warehouse minimizes data replication, improves network bandwidth use and streamlines analytical processing tasks.

- **Improve accuracy and achieve better outcomes using more data points and sophisticated analytical models.** The massively parallel architecture offered by data warehouses is useful for processing larger, more complex information sets. Modelers can easily add new sets of variables if model performance degrades or changes are needed for business reasons.
- **Achieve faster time to results by building, updating and deploying models more quickly.** SAS Analytics Accelerator for Teradata enables analytical processing to be pushed down to the database or data warehouse, shortening the time needed to build and deploy predictive models. It also reduces the latency and complexity associated with the model development process. Analytic professionals have fast access to up-to-date, consistent data and increased processing power. This delivers faster time to results and provides better insights for improved business decision making.
- **Enhance productivity of analytic teams.** In-database analytics helps modelers, data miners and analysts focus on developing high-value analytical modeling tasks instead of spending time consolidating and preparing data.



## Product Overview

SAS Analytics Accelerator for Teradata provides a core set of statistical and analytical functions for in-database processing. The critical data-intensive computations for each SAS analytical procedure are moved inside the Teradata database as vendor-defined functions. Sampling and binning macros from SAS Enterprise Miner execute SQL processing in the database.

The included SAS procedures are commonly used to execute the following model development steps inside the database:

- Optional sampling (e.g., oversampling, rare target events and stratified sampling).
- Data summarization and discovery.
- Time series data aggregation, including trend and seasonal analysis.
- Binning interval predictors.
- Variable selection and dimension reduction.
- Predictive modeling (e.g., multiple linear and logistic regression)
- Scoring of linear models.

In addition, there are three options for specifying which type of in-database computing should be performed: none, with specific SAS procedures or whenever possible.

SAS offers several high-performance computing options for processing large, complex sets of data using high-powered analytics, taking into consideration the needs of both IT and business users to provide the right amount of computing power to solve unique customer issues.

## Key Features

### Statistical and analytical functions enabled for in-database processing

- SAS/STAT® in-database versions of the following procedures:
  - CORR (correlation).
  - CANCELL (canonical correlation).
  - FACTOR (factor).
  - PRINCOMP (principal components).
  - REG (regression analysis, including stepwise regression).
  - SCORE (scoring of linear models).
  - VARCLUS (group variables into clusters).
- SAS/ETS® in-database version of the following procedure:
  - TIMESERIES (analyzes time-stamped transactional data and aggregates the data into a time series format for trending and seasonal analysis).
- SAS® Enterprise Miner™ in-database versions of the following macros and procedures:
  - Sampling macro (simple random, stratified and oversampling).
  - Binning macro (quantile binning with Weights of Evidence calculation).
  - DMDB (data mining database model effects summarization).
  - DMINE (variable selection).
  - DMREG (linear and logistic regression analysis, including support for stepwise regression).

### SQL generation options specify the type of in-database computing

- NONE (specifies that no in-database computation be performed).
- DBMS (specifies the SAS procedure to be used for in-database processing when possible):
  - Uses conventional SAS processing when the specific procedure statement and options do not support in-database processing.
- ALL (specifies that in-database computation be performed whenever possible).
- These options can be specified as either a LIBNAME statement option or as a system option in an OPTIONS statement.

### SAS® Analytics Accelerator for Teradata System Requirements

To learn more about SAS Analytics Accelerator for Teradata system requirements, download white papers, view screenshots and see other related material, please visit [www.sas.com/analyticsaccelerator](http://www.sas.com/analyticsaccelerator).