Optimizing Your Analytics Life Cycle with SAS & Teradata

Rick Lower
Agenda

• The Analytic Life Cycle

• Common Problems

• SAS & Teradata solutions
Analytical Life Cycle

- **Exploration**: Explore All Your Data
- **Preparation**: Prepare Data for Analytics
- **Development**: Build Analytic Models
- **Deployment**: Deliver Results to Business
Analytical Life Cycle

Stage 1
- what the data looks like
- what variables are in the data set
- whether there are any missing observations
- how are the data related
- what are some of the data patterns

Stage 2
- combining data from numerous sources
- handling inconsistent or non-standardized data
- cleaning dirty data
- integrating data that was manually entered
- dealing with semi-structured and structured data

Stage 3
- customer retention
- customer attrition/churn
- marketing response
- consumer loyalty and offers
- fraud detection
- credit scoring
- risk management

Stage 4
- the probability of responding to a particular promotional offer
- the risk of an applicant defaulting on a loan
- the propensity to pay off a debt
- the likelihood a customer leave/churn
- the probability to buy a product
Common Problems
Typical Customer Challenges

• “My analytical process runs too slowly*”
• “We spend too much time moving data around between systems”
• “I want to make more/better use of my EDW/data platform”
• “It takes forever to extract and score the data”
• “There is too much data for us to analyse”
• “We can’t buy new hardware”
• “The quality of our analytical models is lower because we sample – I worry we are missing valuable segments”

* scoring / analysis / data quality / data transformation
The Analytic Data Warehouse
Data Management: The crux of the issue facing your analysts

BUSINESS PROBLEM

Preparing to solve the problem

80%

BUSINESS DECISION

Solving the problem

20%
Data Management: SAS & Teradata working to change the Equation

BUSINESS PROBLEM

Preparing to solve the problem

Solving the problem

BUSINESS DECISION

20% 80%
Barriers to the Adoption of Analytics

- **Scarcity of analytical skills**
  *The need to grow analytical talent from within*

- **Tools that aren’t right for the job**
  *Learning curve to create, share and collaborate*

- **Disjointed, inefficient workflow**
  *How can you fail fast & learn to refine quickly*
SAS & Teradata Solutions
Analytical Life Cycle

**Preparation**
- ACCESS to Teradata
- Code Accelerator
- Data Quality Accelerator
- Data Set Builder for SAS

**Exploration**
- Visual Analytics & Visual Statistics
- Teradata Appliance for SAS

**Development**
- Analytics Accelerator for Teradata
- SAS High-Performance Analytics Products
- TD Appliance for SAS

**Deployment**
- SAS Scoring Accelerator
- SAS Model Manager
- Teradata Appliance for SAS
SAS Analytics in Teradata

- Information Management
- Data Mining / Anomaly Detection
- Text Analytics
- Predictive Analytics
- High Performance Analytics
- Fraud & Risk Detection
- Reporting & Visualization

ARCHITECTURE FLEXIBILITY: SMP - MPP - HADOOP - GRID - ESP

Integrated End-to-End Foundation
Modeling Architecture

Traditional Architecture

In-Database Architecture

Model Development | Model Deployment

Model Development | Model Deployment

Copyright © 2013, Teradata and SAS Institute Inc. All rights reserved.
In-Teradata Analytical Process

- Model Development
- SAS STAT
- SAS Analytics Accelerator
- Business Understanding
- SAS Enterprise Miner
- Model 720 - "Companion"
- Analytic Data Labs Infrastructure
- Scoring ADS

85% of the Development Process

Competitive Advantage

Time To Intelligence

Analytic Insight

Model Execution

Model Deployment
Model Management

Competitive Advantage

Model Development

Faster Analytic Insights
By Accelerating the Entire Analytics Life Cycle

SAS STAT
SAS Analytics Accelerator

SAS Enterprise Miner

Business Understanding

Model 720 - “Companion”

85% of the Development Process

Time To Intelligence

Copyright © 2012, Teradata and SAS Institute Inc. All rights reserved.
The Teradata Environment is SAS Enabled

Optimizing performance by tightly integrating the data and the analytics

- SAS Viya and SAS 9 can be installed directly within the Teradata Environment
- Provides high-speed connectivity between SAS and the data
- SAS in-Database tools further improve data access and performance
- Available as Cloud, On-Premise and Hybrid Architectures
- Supports a single version of data for all analytics across SAS 9 and SAS Viya
- Fully scalable to grow with the customers needs
Achieving Business Lead Outcomes with SAS

Teradata continues to develop and support collaborative SAS & Teradata solutions

<table>
<thead>
<tr>
<th>Horizontal Offers</th>
<th>Vertical Solutions</th>
<th>Dedicated Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Advanced Analytics for Hadoop</td>
<td>• Credit Risk</td>
<td>Teradata 750 Appliance for SAS</td>
</tr>
<tr>
<td>• Analytic Advantage (in-database analytics)</td>
<td>• Credit Scoring</td>
<td></td>
</tr>
<tr>
<td>• Data Quality Advantage</td>
<td>• AML (Anti-Money Laundering)</td>
<td></td>
</tr>
<tr>
<td>• In-database Decision Management</td>
<td>• SAS Asset Performance Analytics</td>
<td></td>
</tr>
<tr>
<td>• Managed Server for SAS</td>
<td>• SAS Field Quality Analytics</td>
<td></td>
</tr>
<tr>
<td>• SAS MA for Teradata</td>
<td>• SAS Production Quality Analytics</td>
<td></td>
</tr>
<tr>
<td>• Teradata Appliance for SAS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SAS Event Stream Processing (ESP) and TD Listener</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Accelerated Insights</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Optimization Services</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# In-Database Functionality

<table>
<thead>
<tr>
<th>SAS/Access to Teradata:</th>
<th>SAS Code Accelerator for Teradata</th>
<th>SAS Scoring Accelerator for Teradata</th>
</tr>
</thead>
<tbody>
<tr>
<td>• PROC APPEND</td>
<td>• PROC DS2</td>
<td>• EM/STAT* Models</td>
</tr>
<tr>
<td>• PROC CONTENTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC COPY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC DATASETS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC DELETE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC FORMAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC FREQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC MEANS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC PRINT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC RANK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC REPORT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC SORT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC SQL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC SUMMARY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• PROC TABULATE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SAS Analytics Accelerator for Teradata</th>
<th>Sas Scoring Accelerator for Teradata</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical Analysis Procedures:</strong></td>
<td>• PROC SCORE works with coefficients from:</td>
</tr>
<tr>
<td>• PROC CANCORR</td>
<td>• PROC ACECLUS</td>
</tr>
<tr>
<td>• PROC CORR</td>
<td>• PROC CALIS</td>
</tr>
<tr>
<td>• PROC FACTOR</td>
<td>• PROC CANDISC</td>
</tr>
<tr>
<td>• PROC PRINCOMP</td>
<td>• PROC DISCRIM</td>
</tr>
<tr>
<td>• PROC REG</td>
<td>• PROC FACTOR</td>
</tr>
<tr>
<td>• PROC SCORE</td>
<td>• PROC PRINCOMP</td>
</tr>
<tr>
<td>• PROC TIMESERIES</td>
<td>• PROC TCALIS</td>
</tr>
<tr>
<td>• PROC VARCLUS</td>
<td>• PROC ORTHOREGG</td>
</tr>
<tr>
<td></td>
<td>• PROC VARCLUS</td>
</tr>
<tr>
<td></td>
<td>• PROC QUANTREG</td>
</tr>
<tr>
<td></td>
<td>• PROC REG</td>
</tr>
<tr>
<td></td>
<td>• PROC ROBUSTREG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DQ Accelerator for Teradata</th>
<th>Sas Scoring Accelerator for Teradata</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Match code</td>
<td>• PROC SCORE works with coefficients from:</td>
</tr>
<tr>
<td>• Parsing/Casing</td>
<td>• PROC ACECLUS</td>
</tr>
<tr>
<td>• Gender/Pattern/Identification analysis</td>
<td>• PROC CALIS</td>
</tr>
<tr>
<td>• Standardization</td>
<td>• PROC CANDISC</td>
</tr>
<tr>
<td></td>
<td>• PROC DISCRIM</td>
</tr>
<tr>
<td></td>
<td>• PROC FACTOR</td>
</tr>
<tr>
<td></td>
<td>• PROC PRINCOMP</td>
</tr>
<tr>
<td></td>
<td>• PROC TCALIS</td>
</tr>
<tr>
<td></td>
<td>• PROC ORTHOREGG</td>
</tr>
<tr>
<td></td>
<td>• PROC VARCLUS</td>
</tr>
<tr>
<td></td>
<td>• PROC QUANTREG</td>
</tr>
<tr>
<td></td>
<td>• PROC REG</td>
</tr>
<tr>
<td></td>
<td>• PROC ROBUSTREG</td>
</tr>
</tbody>
</table>
In-Database Example: Data Quality

- SAS Data Quality functions ported to operate in-database
  - Invoked as SQL Stored Procedures
  - Processing leverages parallelism of underlying RDBMS
  - Significant performance/throughput benefits
- Supported functions:
  - Matchcode generation
  - Parsing
  - Standardization
  - Casing
  - Pattern analysis
  - Identification analysis
  - Gender analysis
SAS Visual Analytics & Visual Statistics

**SAS Visual Analytics**
- Data exploration and discovery
- Distribution and summary statistics
- Post-model analysis and reporting

**SAS Visual Statistics**
- Build prediction and classification models
- Refine candidate models
- Compare models and generate score code
Teradata Everywhere Supports SAS Viya

SAS Viya is a cloud-enabled, in-memory analytics engine that delivers everything your customer needs for quick, accurate and consistent results. SAS Viya:

- Provides elastic, scalable and fault-tolerant processing
- Effortlessly scales to meet future needs
- Enables faster processing with in-Memory analytics
- A standardized code base that supports programming in SAS and other languages, like Python, R, Java and Lua
- Support for cloud, on-site or hybrid environments. It deploys seamlessly to any infrastructure or application ecosystem... including Teradata
SAS & Teradata Partner Success Stories

**Kaiser Permanente**
- Increased productivity of the analyst community
- Improved analytic processing times from 73x to 657x for different processes
- Enabled a self-service environment for provisioning and management of data
- Provided better business insights into internal and external data
- Maintained data governance

**Roche**
- Accelerate drug development with integrated data & analytics
- Enabled self-service analytics, data exploration & consumption
- Reduced the cost and time to bring treatments to patients
- Decreased analytic processes from days/hours to just minutes
- Integrated SAS Managed Server within a Teradata 2800 system
- SAS Data Integration, Enterprise Guide and Enterprise Miner

**Cleveland Clinic**
- Improved patient treatments to deliver better care paths
- Captured streaming device data
- Expanded DataLabs for market analysis and strategy
- Decreased operational cost
- Teradata Appliance for SAS (5 node), 2800 Appliance (16 node) and (2) Teradata Appliances for Hadoop
- Multiple SAS in-Database & in-Memory technologies
### Customer Success Story

**Large US Retailer**

Leveraging SAS and Teradata for advanced analytics to analyze customer data and loyalty programs with in-database and in-memory technologies

<table>
<thead>
<tr>
<th>Issues</th>
<th>Solutions</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projects involving high-end analytics were placing more and more demand on IT and Business Analytics teams</td>
<td>Replaced desktop tools with SAS server based HPA tools such as Enterprise Miner, Scoring Accelerator, Model Manager, VA &amp; Data Mining</td>
<td>Improved the speed and performance of analytics</td>
</tr>
<tr>
<td>Needed to streamline the analytics workflow</td>
<td>Added Teradata Appliance for SAS for production and development of data models</td>
<td>Streamlined data preparation, evaluation and testing</td>
</tr>
<tr>
<td>Needed a solution that provided scalability and data consistency</td>
<td>Expanded EDW</td>
<td>Added additional analytics capacity without requiring the addition of new headcount</td>
</tr>
<tr>
<td>Process could not support additional analytics requests</td>
<td>Added Data Labs Environment to streamline data modeling</td>
<td>Freed up IT time used for data collection, loading and prep</td>
</tr>
</tbody>
</table>
• YouTube Videos

• Demonstrate the integration of SAS and Teradata
• From data preparation to data model scoring
• In-database and in-memory portfolio
• 8 videos online
Summary

• Minimize the need to move the data
• Faster modeling times (months/weeks to hours/minutes)
• Improve data quality, availability and consistency
• Work with entire data sets, including enabling an end-to-end view of data from across the enterprise
• Free up staff to focus more time on value-adding activities