

Integration Skills Training from SAS and Teradata

Version 15.00.1

Module 0 - Course Introduction

Teradata/SAS Course Overview	0-3
Introductions	0-4
Course Modules & Timings	0-5
Ground Rules	0-6

Module 1 - Connectivity Testing & In-Database Processing

Recommended Best Practices	1-3
SAS/ACCESS Family of Products	1-4
Testing Teradata/SAS Connectivity.....	1-5
Typical SAS/Teradata Configurations	1-6
Classroom SAS Environment	1-7
Your Teradata & SAS Training Environment	1-8
Data Structures Used in Some Exercises	1-9
Class System Check	1-10
SAS Connectivity Check	1-11
Protecting Your Teradata Password.....	1-12
Encoding Your RDBMS Password.....	1-13
When Using a SAS Server	1-14
Example Code: Password Encoding & Using the Encoded Password	1-15
Log: Password Encoding	1-16
Log: Using the Encoded Password	1-17
Instructor-led Exercise – Using PWENCODE	1-18
More on Program Sharing.....	1-19
What is Teradata?.....	1-20
What is SAS?	1-21
Getting the Best of Both Worlds!.....	1-22
Supporting Architecture	1-23
SAS and Teradata Together	1-24
SAS Analytics: In-Database Principles.....	1-26
In-Database Processing	1-27
Does SAS Programming Change?	1-28
How to Listen in on SAS' Conversation with Teradata.....	1-29
SAS Generally Talks to T/D Via SQL.....	1-30
Exercise: Traditional PROC FREQ	1-31

Associated Editor & Log File Windows.....	1-32
Exercise: In-Database PROC FREQ	1-33
Associated Log File Window	1-34
Exercise: PROC MEANS In-Database.....	1-35
Associated Log File	1-36
“Running In-Database”	1-37
Teradata LIBNAME Code Running In-Database.....	1-38
Do All PROC Options Apply?	1-39
Associated SAS Log	1-40
SQL to Calculate the Median	1-41
SQL to Calculate the Media	1-42

Module 2 - Implicit and Explicit SQL Pass-Through

Recommended Practices.....	2-3
SAS/Teradata: Interface Options.....	2-4
Best Practice Study Relevant Documentation	2-5
SAS Connectivity to Teradata	2-6
GUI Pitfall to Avoid: Often, Double-Clicking = Download	2-7
SAS/Enterprise Guide & Teradata (Preview).....	2-8
Implicit SQL Pass-Through: SAS/ACCESS	2-9
Implicit SQL Pass-Through: Steps	2-10
Implicit SQL Pass-Through: Cautions.....	2-11
Implicit SQL Pass-Through: Guidelines	2-12
How to Tell What Functions Are Passed to Teradata.....	2-13
(Partial) Results	2-14
SAS Functions Automatically Passed to Teradata – via Implicit SQL Pass-Through.....	2-15
Enabling Dynamic Function Mapping.....	2-16
SQL_FUNCTIONS – Used to Extend Translations.....	2-17
Teradata LIBNAME Code Running In-Database.....	2-18
Implicit SQL Pass-Through: Guidelines	2-19
Implicit SQL Pass-Through: Another Example	2-20
Explicit SQL Pass-Through: SAS/ACCESS	2-22
Explicit SQL Pass-Through: Example.....	2-23
Code Submitted Subsequently.....	2-24
Teradata Explain.....	2-25
Explicit SQL Pass-Through: SAS/ACCESS	2-26
Example	2-27
Another Example.....	2-28
Summary: The Two Standard Templates for Teradata SQL Submission from SAS	2-29
Practice Exercises – Using Base SAS	2-30
DBCONDITION: Key Data Set Option	2-31
Exercise.....	2-32
One Solution	2-33
Log File Excerpt	2-34
What if DBCONDITION had been used...?	2-35
LOG File Excerpt	2-36

SAS Output Window	2-37
Reason for This Behavior	2-38
Other Ways to Code This.....	2-39
Solutions.....	2-40
Associated Log File Excerpts	2-41
Options to Predict In-Database Results.....	2-43
Example	2-44
SAS Log Excerpt	2-45

Module 3 - Teradata Concepts Every SAS User Should Know

Basic Teradata Terms.....	3-4
Bringing Basic Terms to Life.....	3-5
(Relevant) Teradata Database Objects	3-6
User Access Rights	3-7
User Defined Functions	3-8
Example Scoring UDF	3-9
User Defined Functions	3-10
SQL_FUNCTIONS & Teradata UDF's.....	3-11
Understanding SPOOL Space	3-12
SQL to Monitor Spool.....	3-13
Example (No Other Query Running).....	3-14
Teradata Viewpoint: Best Way to Monitor/Manage Queries	3-15
Viewpoint: Aborting a Query/Session	3-16
What Makes Teradata So Fast?.....	3-17
Teradata Architecture – Nodes.....	3-18
Teradata Architecture – BYNET	3-19
Teradata Architecture as it Applies to a Typical Business Question.....	3-20
One Reason for Teradata's Speed.....	3-21
Example Data Definition Language (DDL).....	3-22
Primary Keys and Primary Indexes Conceptually Different.....	3-23
How Does the Teradata Database Distribute Rows?	3-24
Primary Indexes	3-25
Creating a Primary Index	3-26
Primary Index Values.....	3-27
Accessing via a Unique Primary Index	3-28
Accessing via a Non-Unique Primary Index.....	3-29
Row Distribution Using a UPI.....	3-30
Row Distribution Using a NUPI	3-31
Row Distribution Using a Highly Non-Unique Index	3-32
Skewed Data	3-33
Access Using Partitioned Data.....	3-34
Logical Example of NPPI vs. PPI.....	3-35
Summary of PPI Benefits.....	3-36
(More) Teradata Indexes	3-37
Accessing Data without Indexes	3-38
SQL That Results in Full Table Scans	3-39

Sampling.....	3-40
Workflow Suggestions.....	3-41
Lesson: the Teradata Sampling Function	3-42
How Teradata Sampling Works.....	3-43
Sampling using the Sample Clause.....	3-44
Example Code.....	3-46
Other Sampling Notes or Considerations	3-47
SAS Data Set: Sample Creation	3-48
Associated SAS Log Excerpt.....	3-49
Exercise.....	3-50
Associated Log File	3-51
Temporary Tables.....	3-52
Suppose You Wanted to Build	3-53
SQL to Create Such a Table	3-54
SQL to Populate This Table	3-55
Submitting from SAS	3-56
Selected Volatile Table Features	3-57
Volatile Table Syntax	3-59
CREATE TABLE DDL SYNTAX	3-60
Simplest DDL Possible?.....	3-61
Column Data Type Attributes.....	3-62
Table Creation from Existing Tables.....	3-63
Example – CREATE TABLE AS Syntax	3-64
Better Example – Because Primary Index Explicitly Declared.....	3-66
Connection = GLOBAL Option	3-67
Connect Options: Mode and Case Sensitivity	3-68
How to Force Casespecific Comparisons in Teradata Mode.....	3-69
Recurrent Theme	3-70
How Should You Split Processing?.....	3-71
Similarities and Differences	3-72
Teradata and SAS – Similar Concepts.....	3-73
Teradata and SAS – Naming Conventions	3-74
Teradata and SAS – Data Types.....	3-75
SAS to SQL Translation Table	3-76
SAS to SQL Basics: Operators & Functions	3-77
SAS to SQL Basics: If-Then-Else	3-78

Module 4 - Reading from & Writing to Teradata from SAS

Extracting Data from Teradata to SAS	4-4
Potential Uses for Teradata Source Tables	4-5
Default Extract Behavior and Options	4-6
Consideration 1: How Large is the Volume of Data?	4-7
Consideration 2: How often will the Data be Accessed by SAS?	4-8
Consideration 3: For What Purpose is the Data Being Accessed?	4-9
Coding Tips to Reduce Data Extract Size	4-10
Module 4: Lab Exercise #1	4-11
Module 4: Lab Exercise #2	4-12
Creating a SAS Data Set	4-13
Creating a SAS Data Set Method 1: Using a SAS Data Step	4-14
Creating a SAS Data Set Method 2: Using PROC SQL	4-17
Creating a SAS Data Set Method 3: Using FastExport	4-18
Revisiting an Earlier Example	4-19
Watch What Happens	4-20
Compare the Logs	4-21
Pitfall to Avoid: Load/Unload Utility Options on LIBNAMES	4-22
Creating a SAS Data Set Method 3: Using TPT FastExport	4-23
Creating a Text File from a Teradata Table – Using a SAS NULL Data Set	4-24
Writing Data to Teradata from SAS	4-25
Ways to Write Data to Teradata Tables	4-26
Key Point to Remember!	4-27
Temporary Table Handling	4-28
DBMSTEMP: SAS 9.2 Option	4-29
How to Create Such a Table, with Data	4-30
Associated Log File	4-31
Creating a More Optimal Table	4-32
Many SAS Syntax Forms to Drop a Table	4-33
Another Way to Create a Volatile Table	4-34
To Append Data from SAS	4-35
Associated Log File	4-36
Syntax to Delete Rows from a Table	4-37
Associated Log File	4-38
Consider the Following Variation	4-39
DBCOMMIT SAS Load Option	4-40
MULTISTMT SAS Load Option	4-41
Uploading/Writing a Permanent Table with FastLoad –SAS DATA Step Syntax	4-42
Uploading/Writing a Table with FastLoad – Using SAS PROC SQL Syntax	4-43
Example Syntax for Uploading acct_nbr	4-44
Note Resultant DDL	4-45
Module 4: Lab Exercise #3	4-46
One Solution	4-47
Associated SAS Log	4-48
Uploading/Writing to a New Table with the FASTLOAD Option, SAS PROC SQL Syntax ..	4-49
Another Solution to the Exercise	4-50
Associated SAS Log	4-51

Joins	4-52
Explicit SQL Pass-Through.....	4-53
Associated SAS Log	4-54
Remember: GUI Tools Won't See Temporary Tables	4-55
Ways to Prove It's There	4-56
Merges and Joins	4-57
Useful SUGI Paper	4-58
Matching SAS & Teradata Concepts.....	4-59
Demonstration.....	4-61
PROC COMPARE.....	4-62
Resulting Output.....	4-63
Teradata-SAS Merge Behavior.....	4-64
Sorting Options	4-65
Ordering Tables and Data Sets Two MERGE Code Samples	4-66
Module 4: Lab Exercise #4.....	4-67
Module 4: Lab Exercise #5.....	4-68
SAS/ACCESS to Teradata – Data Set and Libname Options	4-69
SAS/ACCESS to Teradata – Data Set Options	4-70
Module 4: Independent Lab Exercise #1	4-71
Module 4: Independent Lab Exercise #2	4-72
Module 4: Independent Lab Exercise #3	4-73
Module 4: Independent Lab Exercise #4.....	4-74
Module 4: Independent Lab Exercise #5	4-75
Module 4: Independent Lab Exercise #6	4-76
Module 4: Review Questions.....	4-77

Module 5 - Advanced SQL Structures SAS Can Exploit

Teradata Date Arithmetic.....	5-4
Sample Article.....	5-5
SYS_CALENDAR.CALENDAR.....	5-6
Example: Replacing Use of INTNX	5-7
BIGINT Handling (SAS Problem Note 39831).....	5-8
Exercise: BIGINT Handling	5-9
BIGINT Exercise	5-10
Teradata Subqueries.....	5-12
Code	5-13
Results	5-14
Derived Tables	5-15
Results	5-16
Consider These Four Queries.....	5-17
Results	5-18
Correlated Subqueries	5-19
One Solution	5-20
Results	5-21
Understanding Teradata Analytics.....	5-22
Ordered (or Windowed) Analytic(al) Functions	5-23
Consider the Following Table - TM_WORKSHOP.fc_transactionsns	5-24
Average Transaction Amount for Each Date.....	5-25
Results	5-26
But What about the Cumulative Average?.....	5-27
Syntax for Ordered Analytic Functions – Moving Averages.....	5-28
Interpreting the Syntax.....	5-29
SQL for the Cumulative Average	5-31
Cumulative Average, from SAS	5-32
Results	5-33
SQL for Cumulative & Moving Averages	5-34
Module 5: Lab Exercise # 1	5-35
Module 5: Lab Exercise # 1 Discussion.....	5-36
Moving Differences	5-37
Module 5: Lab Exercise #2	5-38
Module 5: Lab Exercise #2 Solution.....	5-39
Teradata Ranking Functions	5-40
Parameterizing Ranking Functions	5-41
Module 5: Lab Exercise #3	5-42
Module 5: Lab Exercise #3 Solution.....	5-43
Descriptive Statistics.....	5-44
Module 5: Lab Exercise #4	5-45
Variable Transformation Operations.....	5-46
Dummy Coding and Contrast Coding.....	5-47
Module 5: Lab Exercise #5	5-48
Variable Recoding.....	5-49
Module 5: Lab Exercise #6	5-50
Variable Rescaling	5-51

Module 5: Lab Exercise #7	5-52
Standardization	5-53
Module 5: Lab Exercise #8.....	5-54
Miscellaneous Transformations & Functions.....	5-56
Module 5: Lab Exercise #9.....	5-57
Understanding Denormalization.....	5-59
Module 5: Lab Exercise #10.....	5-60
Module 5: Lab Exercise #10 Solution	5-61

Module 6 - The Teradata Data Set Builder for SAS

Contents	6-3
What is the Teradata Data Set Builder for SAS?.....	6-4
How is this Tool Used?.....	6-5
What is the Data Set Builder for SAS?.....	6-6
Profiling Functions & Features.....	6-7
Data Reorganization/Transformation Functions.....	6-8
Using the Data Set Builder (or Profiler) 7 Basic Steps to Results.....	6-9
Demonstration or Instructor-Led Exercise	6-10
Step 5	6-11
Steps 6 & 7.....	6-12
Variations.....	6-13
For SAS Users	6-14
Duplicate (Address) Checking.....	6-15
Explicit SQL Pass-Through.....	6-17
Example: Embedding Duplicate Check in SAS	6-18
Log File Results	6-19
Frequency Analysis	6-20
Frequency Analysis – Options.....	6-21
Values Analysis	6-22
For SAS Users... ..	6-23
Statistics Analysis.....	6-24
Compare to PROC MEANS (Module 1).....	6-25
Histogram Analysis	6-26
Data Reorganization & Transformation	6-28
Remember Cumulative Average?.....	6-29
Example Variable Creation.....	6-30
Results.....	6-31
Remember Subqueries?	6-32
Two Variable Creations Used.....	6-33
Subquery Embedding.....	6-34
Results.....	6-35
Module 5's Transformations.....	6-36
Remember Exercise #10 – Denormalization?	6-37
Results.....	6-38

Module 7 - SAS/Enterprise Guide

Recommended Practices	7-3
SAS Enterprise Guide	7-4
SAS (the Company) & Enterprise Guide	7-8
Enterprise Guide – Behind the Scenes	7-9
Configure EG Properly	7-10
Turn on Debug Options & Project Logging	7-11
Where Should Options be Input?	7-12
Turn SQLGENERATION On.....	7-13
One-Way Frequencies = PROC FREQ	7-14
Associated Project Log	7-15
Summary Statistics = PROC MEANS	7-16
Turn DBIDIRECTEXEC On	7-17
Query Builder Pass-through Options	7-18
Use EG's PREVIEW Options.....	7-19
Leverage Info Maps & Stored Processes	7-20
Example of Sampling Stored Process	7-21
Pitfalls to Avoid	7-22