

DATA MANAGEMENT USER GROUP

DAVE SMITH, SAS UK



- Section 1: 9.3 technology
 - How to make the most of SAS Data Integration with database technologies
- Section 2: 9.4 technology
 - What's new in Lineage?

SECTION 1

DATA INTEGRATION IN DATABASE



SAS 9.3 HOW TO MAKE THE MOST OF SAS DATA INTEGRATION **TECHNOLOGY** WITH DATABASE TECHNOLOGIES

- New SQL transformations designed to work in database
- How to find the right functions for your database
- How to check processing will occur in database
- Database specific loaders

- Create Table
- Extract
- Set Operators
- Insert Rows
- Delete
- Update
- Merge
- Execute

- In-database doesn't allow the use of SAS lookup etc!
- Join Methods:
 - Sort-Merge Joins
 - Most often selected by the SQL optimiser
 - Use the **Suggest Sort Merge Join** property on the SQL Properties pane to encourage a sort-merge
 - Index Joins
 - Looks up each row of the smaller table by querying an index of the large table
 - Use IDXWHERE=YES as a data set option
 - Turn on the **Suggest Index Join** property on the properties panes for the tables in the query
 - Hash Joins
 - The optimizer considers a hash join when an index join is eliminated as a possibility
 - Loads smaller table into memory; selected based upon BUFFERSIZE option

DATA INTEGRATION IN DATABASE

FURTHER OPTIMISATION

- Bulk Load tables
 - Bulkload to DBMS (affects output table)
 - Use Bulkload for Uploading (applies to the source table)
- Optimize the SELECT Statement
 - Do you really need select * ?
 - Non-character columns first, character columns last
- Don't use Buffering options with in-database!
 - Use for SAS query processing, but not in-database
 - Can prevent implicit pass-through
- Use Threaded Reads
 - Partitions the reading of the data from the DBMS
 - Can specify Where clauses for partitioning with DBSLICE

DATA INTEGRATION IN DATABASE IN DATABASE FUNCTIONS

- Use the expression builder
- Discover UDFs

DATA INTEGRATION IN DATABASE

DATABASE SPECIFIC LOADERS

BREAK



- The new Relationship service
- The Lineage viewer
- Relationship between lineage and governance (Business Data Network)

- The Relationship Service collects and stores metadata
 - Content from SAS and sources outside of SAS
 - Processes that include resources used in data management, business intelligence, and data integration
- Consists of Resources and Relationships
 - Resources are metadata representations of data assets or processes
 - Relationships describe how two Resources are related

- Is dependent on
- Is parent of
- Contains
- Is synonymous with
- Is associated with
- Is equal to

- Scheduled Collection and Loading
 - Set up in the SAS Management Console
- On Demand Collection and Loading
 - <SASHOME>\SASPlatformObjectFramework\9.4\tools\admin\.
 - sas-relationship-loader [options...]
- Batch job loading from SAS DM platform
 - pjExportLineage

- sas-relationship-reporter [options...]
- Creates list format or csv for further reporting
- Pre-defined reports are:
 - Lineage – returns resources related to the subject using *is dependent on* and *contains* relationships.
 - Impact - returns resources that have an *is dependent on* relationship from the subject.
 - DirectDependencies - returns the objects for which the subject directly *is dependent on*.
 - IndirectDependencies - returns the objects for which the subject indirectly *is dependent on*.

RELATIONSHIP REPORTER

DEFAULT LINEAGE REPORT

"/Shared Data/Reports/ClassReport.srx" (SAS report)

--> Is dependent on: "/Shared Data/Maps/Class Relational Map" (Information map (relational))

--> Is dependent on: "/Shared Data/SASHELP/CLASS" (Table)

--> Is dependent on: "/Shared Data/SASHELP/SASHELP" (Library)

--> Contains: "Age" (Column)

--> Contains: "Height" (Column)

--> Contains: "Name" (Column)

--> Contains: "Sex" (Column)

--> Contains: "Weight" (Column)

- --> Is dependent on: "/Shared Data/SASHELP/SASHELP" (Library)

--> Contains: "Age" (InformationMapDataItem)

--> Is dependent on: "Age" (Column) - in "/Shared Data/SASHELP/CLASS" (Table)

--> Contains: "Height" (InformationMapDataItem)

--> Is dependent on: "Height" (Column) - in "/Shared Data/SASHELP/CLASS" (Table)

--> Contains: "Name" (InformationMapDataItem)

--> Is dependent on: "Name" (Column) - in "/Shared Data/SASHELP/CLASS" (Table)

--> Contains: "Sex" (InformationMapDataItem)

--> Is dependent on: "Sex" (Column) - in "/Shared Data/SASHELP/CLASS" (Table)

--> Contains: "Weight" (InformationMapDataItem)

--> Is dependent on: "Weight" (Column) - in "/Shared Data/SASHELP/CLASS" (Table)

"/Shared Data/DataSources/CARS" (Table)

- <-- Is dependent on: "/Shared Data/Cubes/BID_MDDBCARS" (Cube)

- <-- Is dependent on: "/Shared Data/Maps/OLAPMaps/MDDBCARS" (Information map (OLAP)) “

/User Folders/sasdemo/My Folder/CARS" (Table)

- <-- Is dependent on: "Parameters" (Prompt group)

- <-- Contains: "/User Folders/sasdemo/My Folder/cars-pak" (Stored process)

- <-- Is dependent on: "/User Folders/sasdemo/My Folder/Cars Report.srx" (SAS report)

- <-- Is dependent on: "Parameters" (Prompt group)

- <-- Contains: "/User Folders/sasdemo/My Folder/cars-stream" (Stored process)

- `sas-relationship-reporter -report lineage -types Report -name "ClassReport.srx" -format list -relContent {-name "CLASS"}`

`"/Shared Data/Reports/ClassReport.srx"` (SAS report)

--> Is dependent on: `"/Shared Data/Maps/Class Relational Map"`
(Information map (relational))

--> Is dependent on: `"/Shared Data/SASHELP/CLASS"` (Table)

- `sas-relationship-reporter -report indirectDependencies -types Report -name "ClassReport.srx" -relContent {-name "AGE"} -format list`

`"/Shared Data/Reports/ClassReport.srx"` (SAS report)

--> Is dependent on: `"/Shared Data/Maps/Class Relational Map"`
(Information map (relational))

--> Contains: "Age" (InformationMapDataItem)

--> Is dependent on: "Age" (Column) - in `"/Shared Data/SASHELP/CLASS"` (Table)

RELATIONSHIP REPORTER

CSV REPORT

- `sas-relationship-reporter -report lineage -types Report -name "ClassReport.srx" -format csv -file ClassReportLineage.csv`

```
ClassReportLineage.csv - Notepad
File Edit Format View Help
SubjectIndex,Depth,LeftName,LeftPath,LeftType,LeftTypeLabel,RelationshipType,Direction,RightName,RightPath,
1,0,ClassReport.srx,/Shared Data/Reports/ClassReport.srx,Report,SAS report,Is dependent on,TO,Class Relatio
1,1,Class Relational Map,/Shared Data/Maps/Class Relational Map,InformationMap.Relational,Information map (I
1,2,CLASS,/Shared Data/SASHELP/CLASS,Table,Table,Is dependent on,TO,SASHELP,/Shared Data/SASHELP/SASHELP,Lil
1,2,CLASS,/Shared Data/SASHELP/CLASS,Table,Table,Contains,TO,Age,Age,Column,Column
1,2,CLASS,/Shared Data/SASHELP/CLASS,Table,Table,Contains,TO,Height,Height,Column,Column
1,2,CLASS,/Shared Data/SASHELP/CLASS,Table,Table,Contains,TO,Name,Name,Column,Column
1,2,CLASS,/Shared Data/SASHELP/CLASS,Table,Table,Contains,TO,Sex,Sex,Column,Column
1,2,CLASS,/Shared Data/SASHELP/CLASS,Table,Table,Contains,TO,weight,weight,Column,Column
1,1,Class Relational Map,/Shared Data/Maps/Class Relational Map,InformationMap.Relational,Information map (I
1,1,Class Relational Map,/Shared Data/Maps/Class Relational Map,InformationMap.Relational,Information map (I
1,2,Age,Age,InformationMapDataItem,InformationMapDataItem,Is dependent on,TO,Age,Age,Column,Column
1,1,Class Relational Map,/Shared Data/Maps/Class Relational Map,InformationMap.Relational,Information map (I
1,2,Height,Height,InformationMapDataItem,InformationMapDataItem,Is dependent on,TO,Height,Height,Column,Coli
1,1,Class Relational Map,/Shared Data/Maps/Class Relational Map,InformationMap.Relational,Information map (I
1,2,Name,Name,InformationMapDataItem,InformationMapDataItem,Is dependent on,TO,Name,Name,Column,Column
1,1,Class Relational Map,/Shared Data/Maps/Class Relational Map,InformationMap.Relational,Information map (I
1,2,Sex,Sex,InformationMapDataItem,InformationMapDataItem,Is dependent on,TO,Sex,Sex,Column,Column
1,1,Class Relational Map,/Shared Data/Maps/Class Relational Map,InformationMap.Relational,Information map (I
1,2,weight,weight,InformationMapDataItem,InformationMapDataItem,Is dependent on,TO,weight,weight,Column,Coli
```

RELATIONSHIP SERVICE

LINEAGE VIEWER

- Administering the tools
- <http://support.sas.com/documentation/cdl/en/bisag/67481/HTML/default/viewer.htm#p1dcleibglhr6wn19yzvfz81yw5y.htm>
- SGF paper
- **Paper SAS396-2014**
- **Understanding Change in the Enterprise**
- Liz McIntosh, Nancy Rausch, and Bryan Wolfe; SAS Institute Inc., Cary, NC



**THE
POWER
TO KNOW[®]**