

## Proof of Concept for ETL Vendors

### SAS Response

Course: Evaluating ETL Tools and Technologies, afternoon session  
ETL Vendors in Action

# Evaluating ETL Tools and Technologies

## Table of Contents

Proof of Concept Overview .....	3
Scenario Overview .....	3
Demo Scenarios / Topics .....	4
SAS® Enterprise Data Integration Server .....	5
1. Company Overview and View of ETL and Data Integration .....	5
2. Installation and setup .....	7
3. Data exploration / profiling.....	14
4. Complex mapping.....	17
5. Maintenance features .....	20
6. Performance .....	26
7. Scheduling and monitoring .....	28

# Evaluating ETL Tools and Technologies

## Proof of Concept Overview

The scenarios for the proof of concept are all based on a wholesale business that supplies specialty products to retailers. The scenarios are based on the items that one might consider important when evaluating an ETL solution for a single data warehouse.

The examples are all built around a wholesale shipments schema, with a set of source tables loaded with data and a set of target tables to be populated by the tools. The extract rules for the schema are simple, but should be enough to demonstrate basic and some advanced capabilities in the products.

The afternoon will be a mix of discussion and demo, with the emphasis on showing how the products are used to accomplish specific tasks. While the focus is on extraction, some of the scenarios or presentation topics involve showing other features like metadata management, data profiling or monitoring job execution.

Because there's no way to show the entire set of ETL for three vendors in the time allotted, we'll be using different elements to show different features. For the scenarios listed we expect to see the features used to accomplish the task live. It isn't expected that we can see the entire extract constructed for each scenario in the time given. However, a complete set of extracts is required in order to show how dependencies, scheduling and monitoring work.

Demo time is limited so there are topics/scenarios labeled "time permitted" which we may not be able to show. They are included in case we have extra time at the end of the class.

## Scenario Overview

In a proof of concept you provide to vendors all the source and target table definitions, extract rules and source data. Since this is meant to reflect the real ETL you'll be doing, it's a good idea to select both simple extracts and complex extracts or extracts that have problem data. When you provide this information, it should be formally documented so the vendor understands in detail what they are supposed to show.

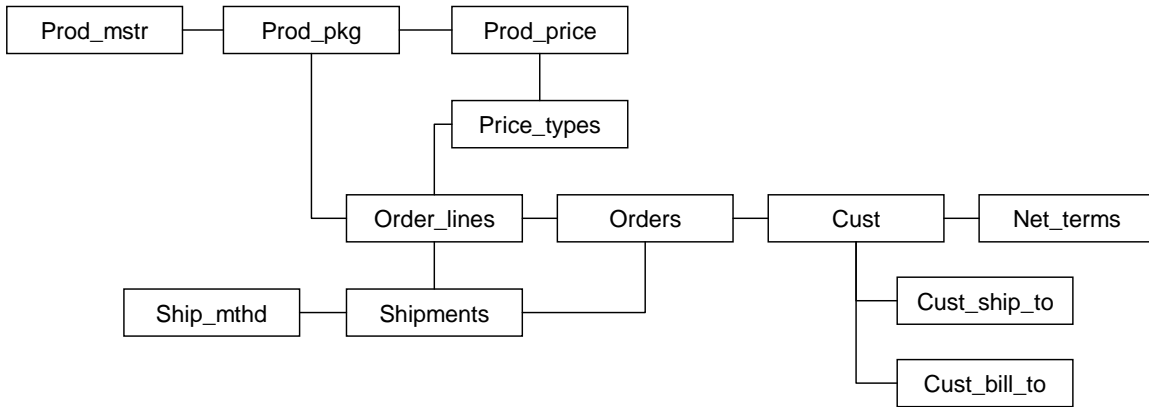
Part of the reason for selecting source data with quality problems is that this will show how a developer is expected to work within the tool. If all the extracts are based on ideal tables and data, as with standard vendor demos, then you won't see what a developer really has to face when dealing with data exceptions.

As a rule, you should have imperfect data, tables with relationship problems like different types on join or lookup columns, and you should always require the use of relational database in the proof of concept.

Using a database is important because it will show you how the tool interacts with a database. Many vendor demos you see are based on text files as input and output, which can avoid some of the difficulties when dealing with SQL since all the work is done directly in the ETL engine.

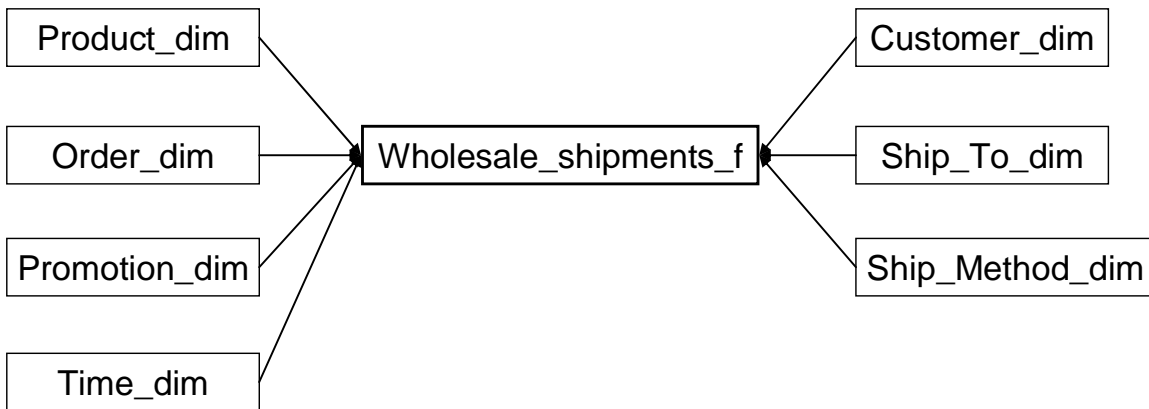
For our scenarios, we will be using the following source and target schemas. The source schema consists of 12 tables with some of the common design problems found in source databases.

## Evaluating ETL Tools and Technologies



Among the problems are a mix of second and third normal form, multi-part keys, invalid foreign key constraints, and multiple join paths to the same data elements. In addition to the above tables there are two change-data-capture tables for the shipment and customer data. These are used for incremental extract examples.

The target schema has seven dimension tables and one fact table. The fact table contains all shipments of order lines by day to customer addresses in the ship\_to dimension.



There are several things in this target schema that complicate the extracts. The time dimension is based on a fiscal calendar using 4 and 5 week periods rather than a simple date-based dimension. There is no source for this data, so it must be constructed by the ETL job. A non-data-driven extract is a challenge for some ETL products. The ship\_method dimension has unique rows based on a multi-part key which can cause trouble for some ETL tools' lookup functions. The specific details about extract rules and data are available at the end of this document.

## Demo Scenarios / Topics

The following section describes each of the presentation topics or scenarios that will be reviewed during the course. For each scenario, there is a list of directions or questions to be answered and a short view of the schema and data (if applicable).

Included with the descriptions are sample criteria you might use during an evaluation so that you can score the vendors during the class. After this section there are responses from each of the vendors to all of the scenario questions so you have something to refer back to.

# SAS® Enterprise Data Integration Server

## 1. Company Overview and View of ETL and Data Integration

- What is your position in the market?

Current studies have indicated that SAS is one of the top three data integration vendors, currently holding nearly 10% of the market. The top of this market is very tight; the other two leading players maintain approximately 13% and 16% market shares respectively. SAS is committed to being a leader in this market space and will continue to increase market share in the area of data integration.

- What do you consider your sweet spot in applications or in industries?

SAS' sweet spot is in helping organizations address difficult business issues for a competitive advantage. SAS is the leader in analytics and Business Intelligence, serving businesses and organizations in all industry segments. These analytics and business intelligence capabilities are fully integrated with SAS Enterprise Data Integration Server, creating a robust platform designed to deliver answers quickly to decision makers throughout the organization.

- What is the core ETL product offering?

The flagship product offering is SAS Enterprise Data Integration Server, which includes the SAS Data Integration Studio, a full-featured design and management environment for ETL and Data Integration. SAS Enterprise Data Integration Server also includes an integrated metadata environment offering impact analysis and data lineage as well as embedded data quality and profiling.

- What are your product's strongest points?

**Eliminates overlapping and redundant tools.** Regardless of the projects an IT department faces – building a data warehouse, mirroring databases, summarizing multiple files into one, consolidating or upgrading ERP systems, augmenting operational data with new information, moving data to and from operational systems like SAP – SAS can handle all data integration needs, both now and in the future.

**Lower project costs, increase staff productivity and speed delivery of results.** SAS offers the only comprehensive enterprise data integration environment that is built from the ground up to meet the full spectrum of enterprise data integration needs. Instead of linking and managing technologies from different vendors, SAS Data Integration provides a collaborative design environment promoting reuse, code sharing, management control, shared metadata, design workflow, a rapid learning curve, and ease of maintenance. SAS' comprehensive and integrated solution results in lower overall costs, reduced risk and faster results.

**Deliver consistent and trusted information throughout the enterprise.** SAS Data Integration provides increased confidence in the accuracy and timeliness of operational data and business

## Evaluating ETL Tools and Technologies

information. Built in data enrichment and data quality components add value to data, increasing confidence, and ensuring the best data possible for operational and decision-support use.

- Why would a customer choose your product over another vendor's?

A customer would choose SAS over other vendors because SAS offers a comprehensive platform that integrates Data Integration, Business Intelligence and Analytics. SAS' complete portfolio of capabilities allows an organization to realize data consistency as well as meet the needs of all users across the organization. SAS offers a complete, reliable and flexible offering that enables an organization to meet short-term goals and objectives while positioning an organization for long-term success.

**SAS is Complete** - SAS has built from the ground up a comprehensive enterprise data integration environment. SAS data integration technology is built entirely by SAS (not purchased from other vendors) so it works seamlessly. Shared metadata, integrated data quality and shared common components guarantee that all SAS products and solutions are tightly integrated.

**SAS is Flexible** - Regardless of the customers' Data Integration needs, from simple, one-time migrations to complex, real-time integrations, the SAS technology can meet those needs – regardless of a customer's unique circumstances. The solution also grows with the customer's needs.

**SAS is Reliable** - SAS pioneered the idea of Data Integration over 30 years ago. The reliability, stability and accuracy of our offerings is validated by our large and loyal customer base, our industry ranking as a leading Data Integration vendor, and a 98% customer retention rate – the highest in the industry.

SAS is also a company customers count on. SAS stands firm as a stable and dependable partner. We are independent, financially strong with over 30 years of profitability and have an industry-leading commitment to R&D, with over 24% of revenue reinvested in R&D annually.

- SAS view of ETL and Data Integration

SAS views ETL as one of the key components of the overall data integration landscape. Those components include:

**Connectivity:** Connectivity to more data sources on more platforms than any other solution; provides the ability to get data from, or put data to, almost anywhere; support for Message-Oriented Middleware; access to static and streaming data.

**Data cleansing and enrichment:** Data quality embedded into batch, near-time and real-time processes; data quality rules are callable through message queues, Web services and custom exits; an array of out-of-the-box standardization rules– or, build customized rules for special situations; metadata built and shared across the entire process; simple process for institutionalizing data quality business rules; profiling and data quality monitoring.

**Extract, transform and load (ETL):** Transformation library with more than 300 predefined transformations; Transformation Generator wizard; transformations can run on any platform with

## Evaluating ETL Tools and Technologies

any data source; transformation and processes can be deployed easily as embedded business logic or Web services for use in other applications or as part of a service-oriented architecture.

**Migration and synchronization:** Metadata-driven access to sources; extensive and shareable library of predefined transformations for migration and synchronization; embedded, reusable data quality business rules clean data as it is moved; real-time data services for synchronization and migration projects; change data management and capture.

**Data federation:** Virtual access to database structures, enterprise applications, mainframe legacy files, text, XML, message queues and a host of other sources; virtual data across these data sources for real-time access, reporting and analysis; query optimization for homogenous and heterogeneous data sources.

**Master data management:** Semantic data descriptions of input and output data sources uniquely identify each instance of a business element (customer, product, account, etc.) and standardize the master data model; reusable business rules clean, standardize, match and enhance data as it moves into the master reference file and is reused for downstream processes; supports client requests via standard SOAP interfaces; data feeds can arrive in a single transaction or hundreds of transactions at a time.

## 2. Installation and setup

- What is installed on the ETL server?

The SAS Enterprise Data Integration Server is SAS' flagship offering in this space, which includes ETL. The product components installed on the server include:

- Foundation SAS – ETL Engine, Application Server, Batch Server, Metadata Server
- SAS/Access technology – read/write access to third-party data sources
- SAS Data Quality
- Platform Suite for SAS – Job scheduling, load balancing, etc.

- What is installed on the developer and/or administrator's PC?

The SAS Enterprise Data Integration Server is SAS' flagship offering in this space, which includes ETL. The product components installed on the developer and/or administrator's PC include:

- SAS Data Integration Studio
- SAS DataFlux dfPower Quality, Profile, Explorer, Customize
- SAS Management Console

- What other installation is required (e.g. a relational database for metadata)?

Additional installation is not required, but is optional depending on the customer configuration. For instance, the SAS Metadata Server can reside in SAS or on a relational database. SAS native access to relational databases require their client libraries to be available on the execution server. All other technologies are integrated into the installation and configuration process and wizard.

# Evaluating ETL Tools and Technologies

- What post-installation tasks are required?

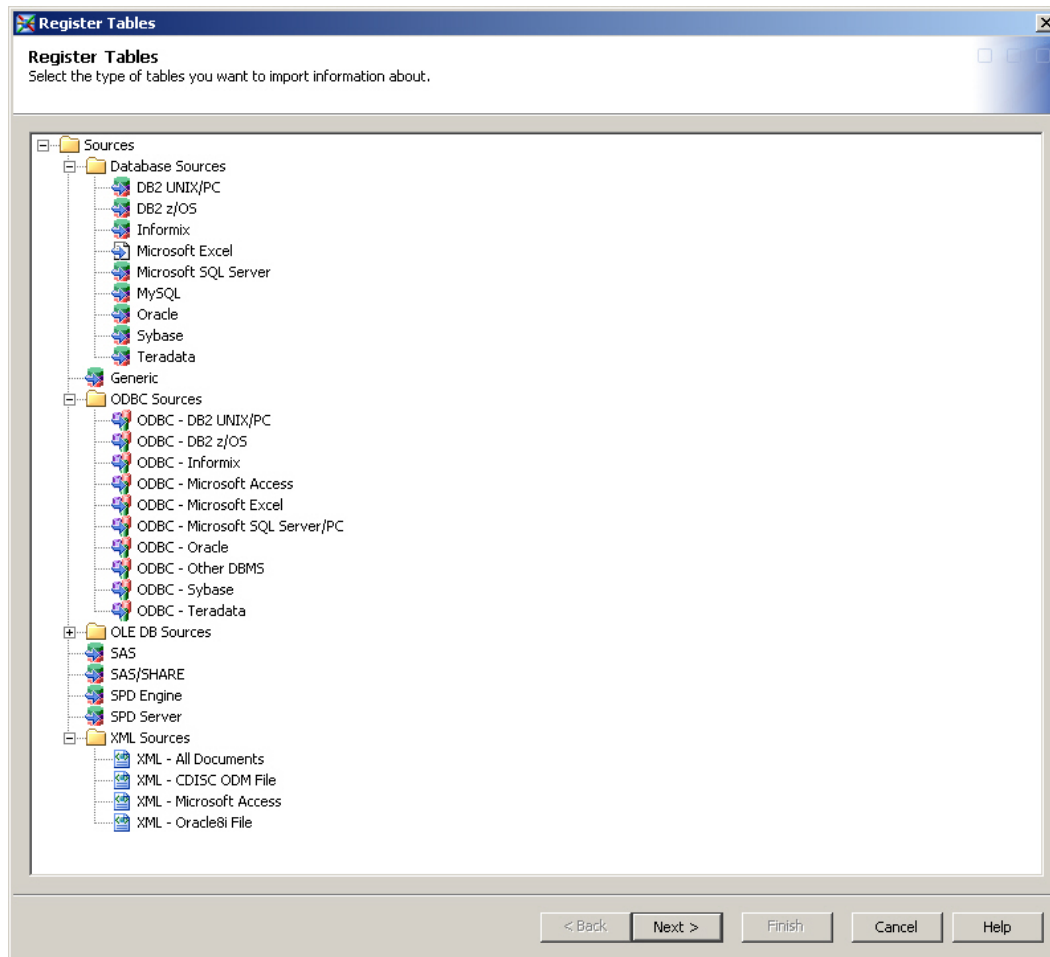
A single configuration step is required. A Configuration Wizard is available to automate the server set-up, authorizations and security.

- How long does it take to install the product to the point where development can start?

Installation and configuration is a function of the complexity of the implementation. For instance, the entire configuration can be deployed on a single homogeneous environment or across multiple heterogeneous environments. SAS Enterprise Data Integration Server can be deployed on an unlimited number of developer/administrator environments. A single machine installation/configuration takes about 30 minutes. In a multiple machine installation/configuration can take 1-3 hours. A Configuration Wizard is available that provides a scripted environment for configuring metadata, servers and user credentials. Manual configuration is also supported which provides greater flexibility, but would require additional time to complete.

## 1. Extractions

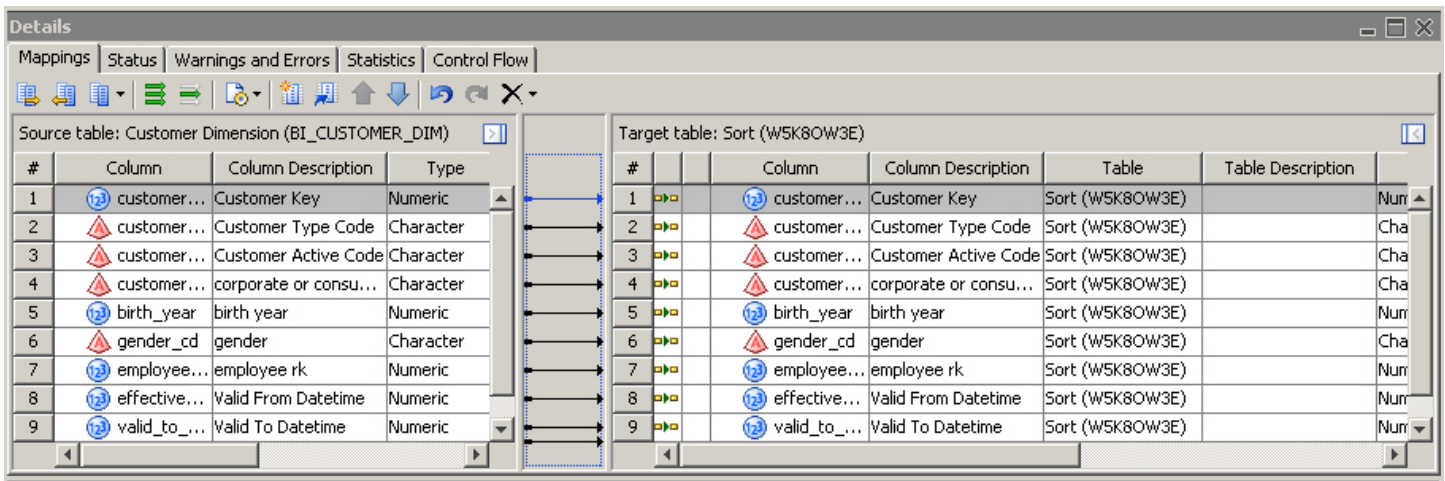
- How do you set up the source and target metadata?



## Evaluating ETL Tools and Technologies

Select the appropriate source or target structure and step through the Wizard to register the appropriate metadata.

- Demonstrate a simple mapping from one source to a target.



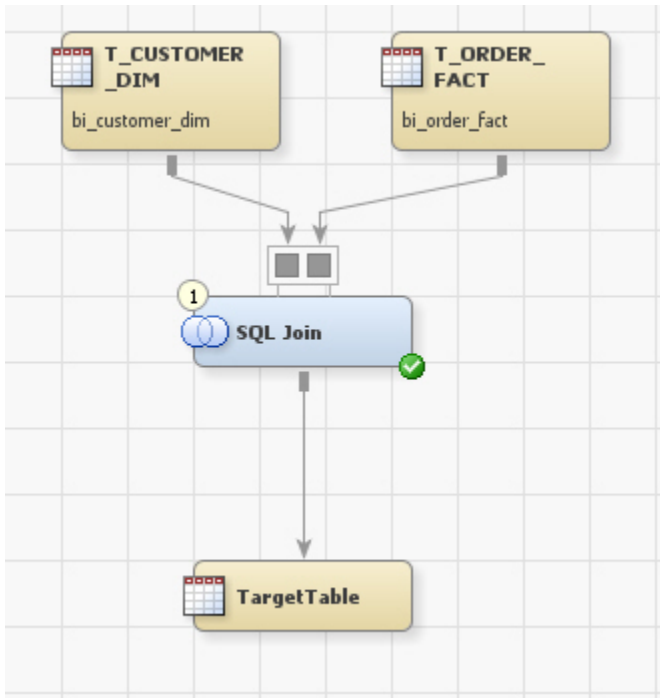
The screenshot shows the 'Details' window of an ETL tool. It displays a mapping between a source table and a target table. The source table is 'Customer Dimension (BI\_CUSTOMER\_DIM)' and the target table is 'Sort (W5K8OW3E)'. The mapping is one-to-one, with each column in the source table mapped to a corresponding column in the target table. The columns are listed in two tables, with arrows indicating the mapping between them.

Source table: Customer Dimension (BI_CUSTOMER_DIM)				Target table: Sort (W5K8OW3E)				
#	Column	Column Description	Type	#	Column	Column Description	Table	Table Description
1	customer...	Customer Key	Numeric	1	customer...	Customer Key	Sort (W5K8OW3E)	Nun
2	customer...	Customer Type Code	Character	2	customer...	Customer Type Code	Sort (W5K8OW3E)	Cha
3	customer...	Customer Active Code	Character	3	customer...	Customer Active Code	Sort (W5K8OW3E)	Cha
4	customer...	corporate or consu...	Character	4	customer...	corporate or consu...	Sort (W5K8OW3E)	Cha
5	birth_year	birth year	Numeric	5	birth_year	birth year	Sort (W5K8OW3E)	Nun
6	gender_cd	gender	Character	6	gender_cd	gender	Sort (W5K8OW3E)	Cha
7	employee...	employee rk	Numeric	7	employee...	employee rk	Sort (W5K8OW3E)	Nun
8	effective...	Valid From Datetime	Numeric	8	effective...	Valid From Datetime	Sort (W5K8OW3E)	Nun
9	valid_to_...	Valid To Datetime	Numeric	9	valid_to_...	Valid To Datetime	Sort (W5K8OW3E)	Nun

Both simple (one to one) and complex (derived) mappings are supported between source and target tables.

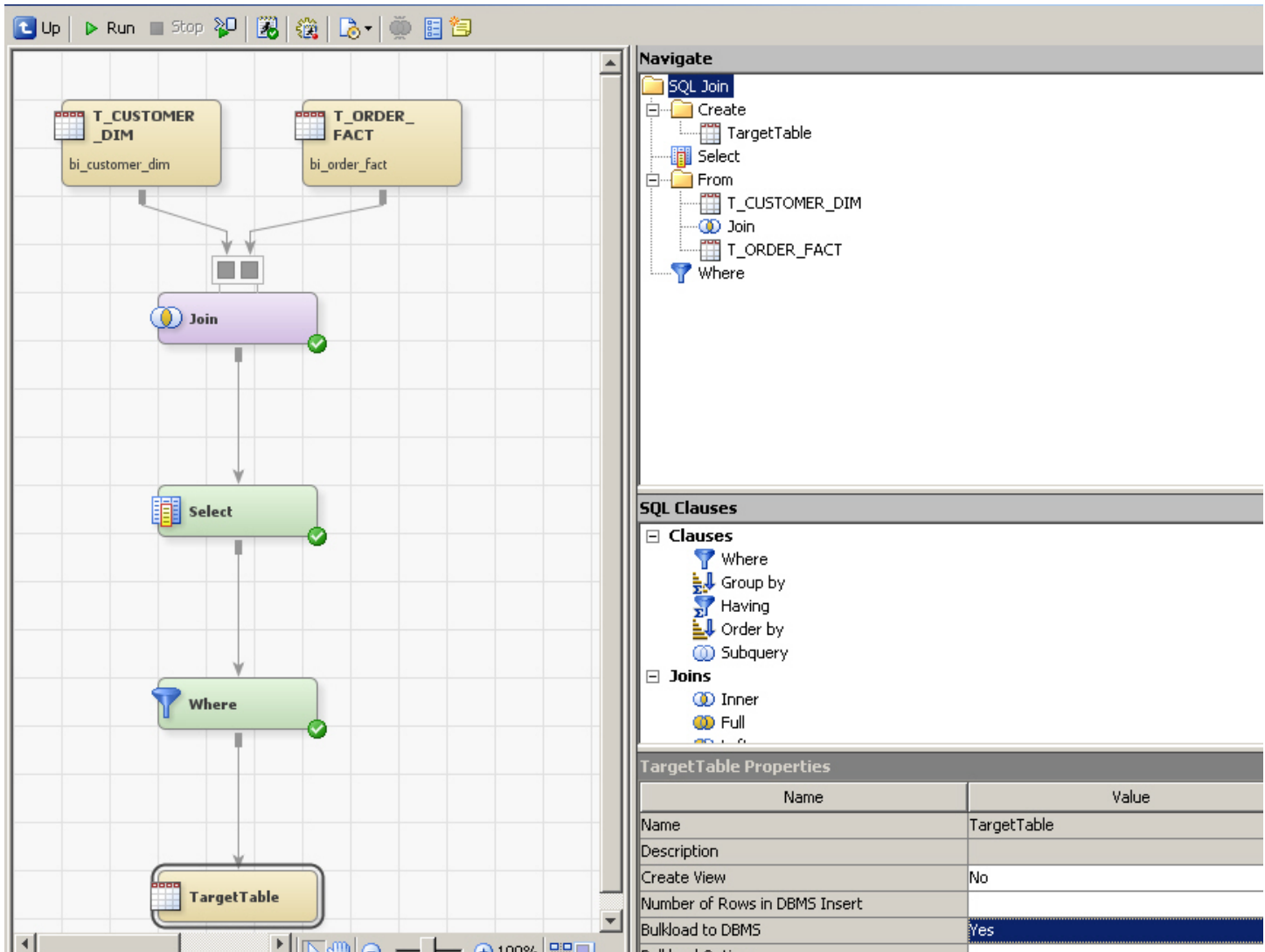
- Demonstrate a simple mapping that joins two tables and populates a single target.

## Evaluating ETL Tools and Technologies



The SQL Transform provides the ability to complete simple SQL extracts and complex SQL joins.

# Evaluating ETL Tools and Technologies



The SQL properties provide the ability to graphically design the SQL join type and join criteria.

- Does the tool automatically map same-name columns between sources and targets?

Details

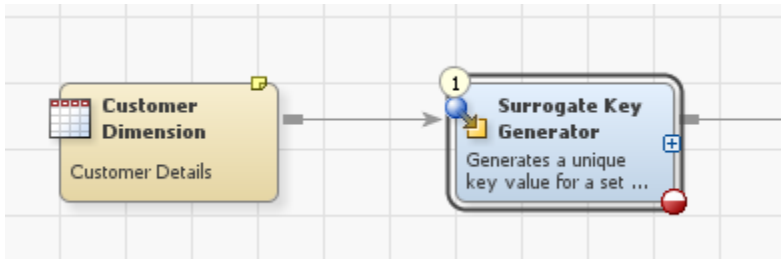
Mappings | Status | Warnings and Errors | Statistics | Control Flow

Source table: Cus[Map all columns (BI_CUSTOMER_DIM)]				Target table: Sort (WSK8OW3E)						
#	Column	Column Description	Type	#	Column	Column Description	Table	Table Description	Type	Length
1	customer...	Customer Key	Numeric	1	customer...	Customer Key	Sort (WSK8OW3E)		Numeric	8
2	customer...	Customer Type Code	Character	2	customer...	Customer Type Code	Sort (WSK8OW3E)		Character	32
3	customer...	Customer Active Code	Character	3	customer...	Customer Active Code	Sort (WSK8OW3E)		Character	32
4	customer...	corporate or consu...	Character	4	customer...	corporate or consu...	Sort (WSK8OW3E)		Character	10
5	birth_year	birth year	Numeric	5	birth_year	birth year	Sort (WSK8OW3E)		Numeric	8
6	gender_cd	gender	Character	6	gender_cd	gender	Sort (WSK8OW3E)		Character	3
7	employee...	employee rk	Numeric	7	employee...	employee rk	Sort (WSK8OW3E)		Numeric	8
8	effective...	Valid From Datetime	Numeric	8	effective...	Valid From Datetime	Sort (WSK8OW3E)		Numeric	8
9	valid_to...	Valid To Datetime	Numeric	9	valid_to...	Valid To Datetime	Sort (WSK8OW3E)		Numeric	8
10	processe...	Processed Datetime	Numeric	10	processe...	Processed Datetime	Sort (WSK8OW3E)		Numeric	8

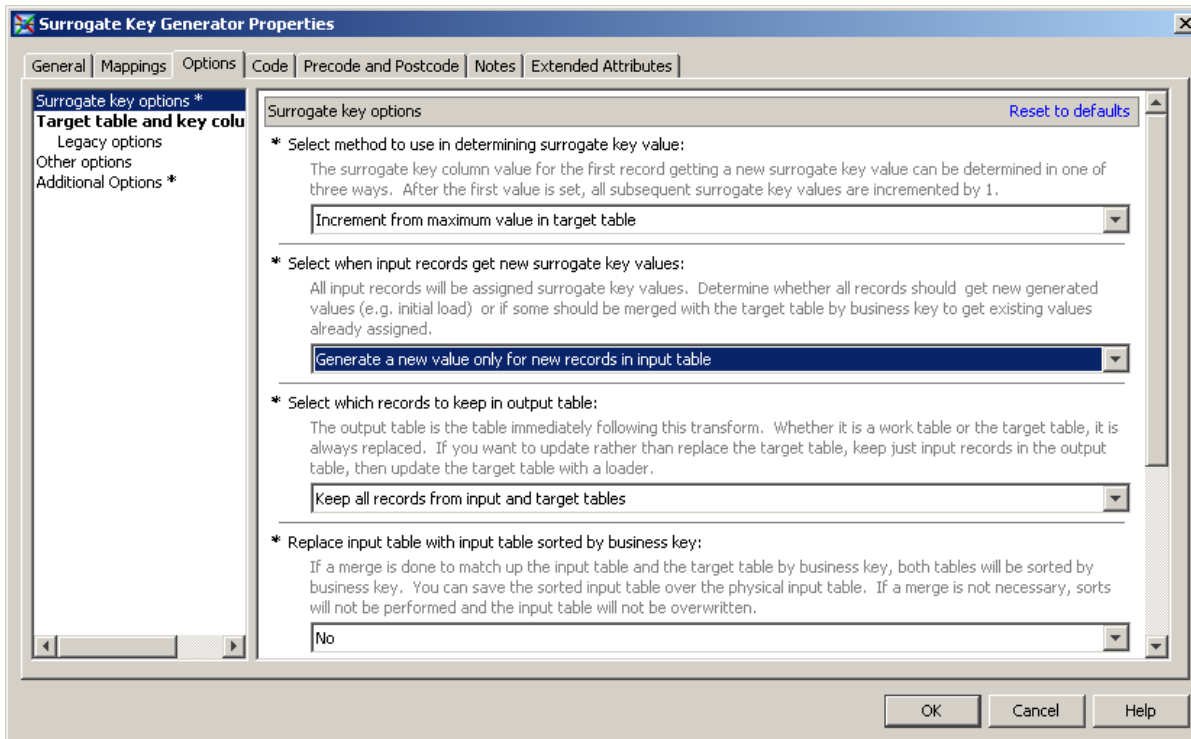
## Evaluating ETL Tools and Technologies

Map All Columns allows the automatic mapping of source/target columns with the same name and attributes.

- How do you suggest storing and maintaining surrogate key values and retaining the mapping of surrogate key values back to production keys?

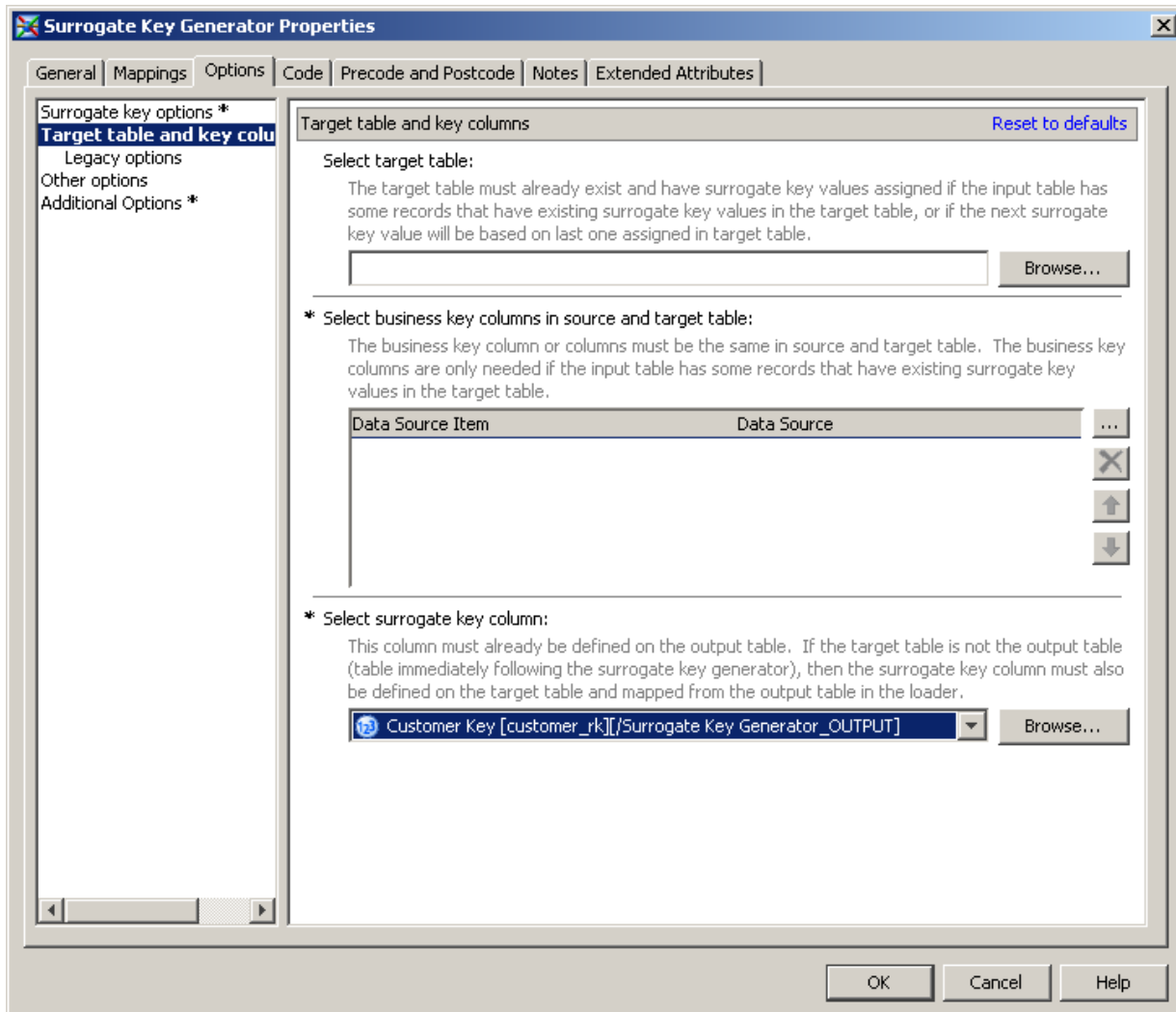


# Evaluating ETL Tools and Technologies



The Surrogate Key Generator transformation produces a surrogate key in the target table.

# Evaluating ETL Tools and Technologies



The Slowly Changing Dimension (SCD) Type 2 Loader allows optimized change detection, selection of columns for change detection, support for business and generated (surrogate) keys.

## 3. Data exploration / profiling

- How do you preview source data from within the design interface?

## Evaluating ETL Tools and Technologies

The screenshot displays a BI tool interface with a main window titled "View Data: Customer Dimension (Filter applied)". The window shows a table with columns: #, customer\_rk, customer\_type\_cd, customer\_active\_cd, customer\_type, birth\_year, and gender. A "Query Options" dialog box is open, showing a filter applied to the "birth\_year" column with the condition "Greater than 1968".

#	customer_rk	customer_type_cd	customer_active_cd	customer_type	birth_year	gender
1	3020	Diamond	Low	corporate	1970	M
2	4020	Emerald	Low	corporate	1970	M
3	5020	Diamond	Low	corporate	1975	F
4	7020	Emerald	Medium	corporate	1975	F
5	9020				1970	F
6	10020				1980	F
7	11020				1970	F
8	13020				1984	M
9	22020				1984	M
10	25020				1984	M
11	26020				1984	M
12	28020				1984	M
13	30020				1984	F
14	34020				1980	M
15	37020				1984	M
16	39020				1980	M
17	42020				1975	M
18	43020				1970	M
19	44020	Emerald	Medium	corporate	1980	M
20	45020	Diamond	High	corporate	1975	F
21	46020	Emerald	High	corporate	1984	F
22	47020	Emerald	Medium	corporate	1980	M
23	49020	Emerald	High	corporate	1980	F

Viewing data is supported on Source, Target and intermediate files in job flows, custom and inventory views of table metadata.

- What data profiling capabilities are available from within the design environment?

There are various capabilities available within the design environment to profile data or assess the type of data and the quality of information. There is a View Statistics capability that provides a detailed overview of tables or views (SAS and DBMS structures). There is also the ability to launch the SAS DataFlux dfPower Profile tool to provide a more graphical representation of the data. This profiling capability provides various metrics to measure the quality of the data, distribution of values, patterns of the data and much more.

# Evaluating ETL Tools and Technologies

The screenshot displays the SAS Data Integration Studio interface. The main window shows the 'Table: prod\_mstr.csv' with 53 total records. A table below lists field statistics:

Field Name	Ordinal Position	Count	Null Count	Blank Count	Minimum Value
cat	5	53	0	0	Beverages
long_desc	4	53	1	0	The Bro, A
msr	2	53	0	(not applicable)	0.5
shrt_desc	3	53	0	0	Albatross
size	7	53	47	0	A cup
upc	1	53	0	(not applicable)	10478516528
wt	6	53	10	0	1 lb

The 'Field: cat' section shows a 'Frequency Distribution' report visualization as a 3D pie chart titled 'cat Frequency Distribution'. The legend includes categories like Beverages, Canned Foods, Beverages - Alcoholic, Carbonated Beverages, Health Foods, Packaged Foods, Breakfast Cereal, Household Appliances, Gifts, Consumer Electronics, Pet Food, and Toys.

- Can you generate rules based on the data profiling and incorporate them within the extracts being built?

Business rules can be created directly within the Profiling tool and then used as a part of an in-stream or embedded process within SAS Data Integration Studio.

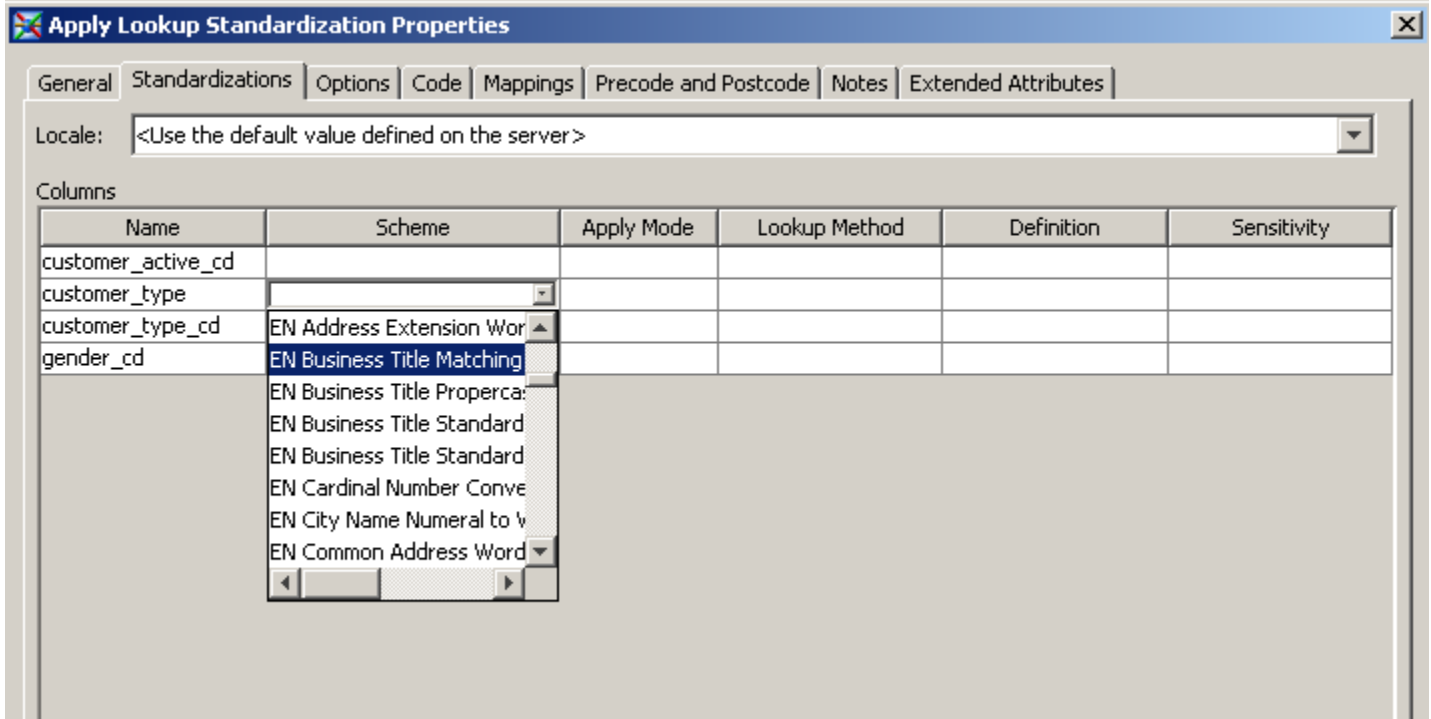
The screenshot shows the 'Scheme Builder' window. On the left, a list of 36 values is shown under the 'Report' tab, including items like 'Albatross', 'Baby Smokes-A-Lot Baby Doll', and 'Choco Dog Chocolate Covered Hot Dogs'. The 'Scheme' tab on the right shows a list of data entries. A 'Save As' dialog box is open, allowing the user to create a new business rule. The dialog shows the following details:

- Item name: Long\_Desc
- Item description: Business Rule for Long Descriptions
- Created by: [blank]
- Creation date: [blank]
- Last modified: [blank]

The dialog also lists various standard rules such as 'EN Address Extension Words', 'EN Business Title Matching', and 'EN Country'.

# Evaluating ETL Tools and Technologies

Create the business rule to standardize values.

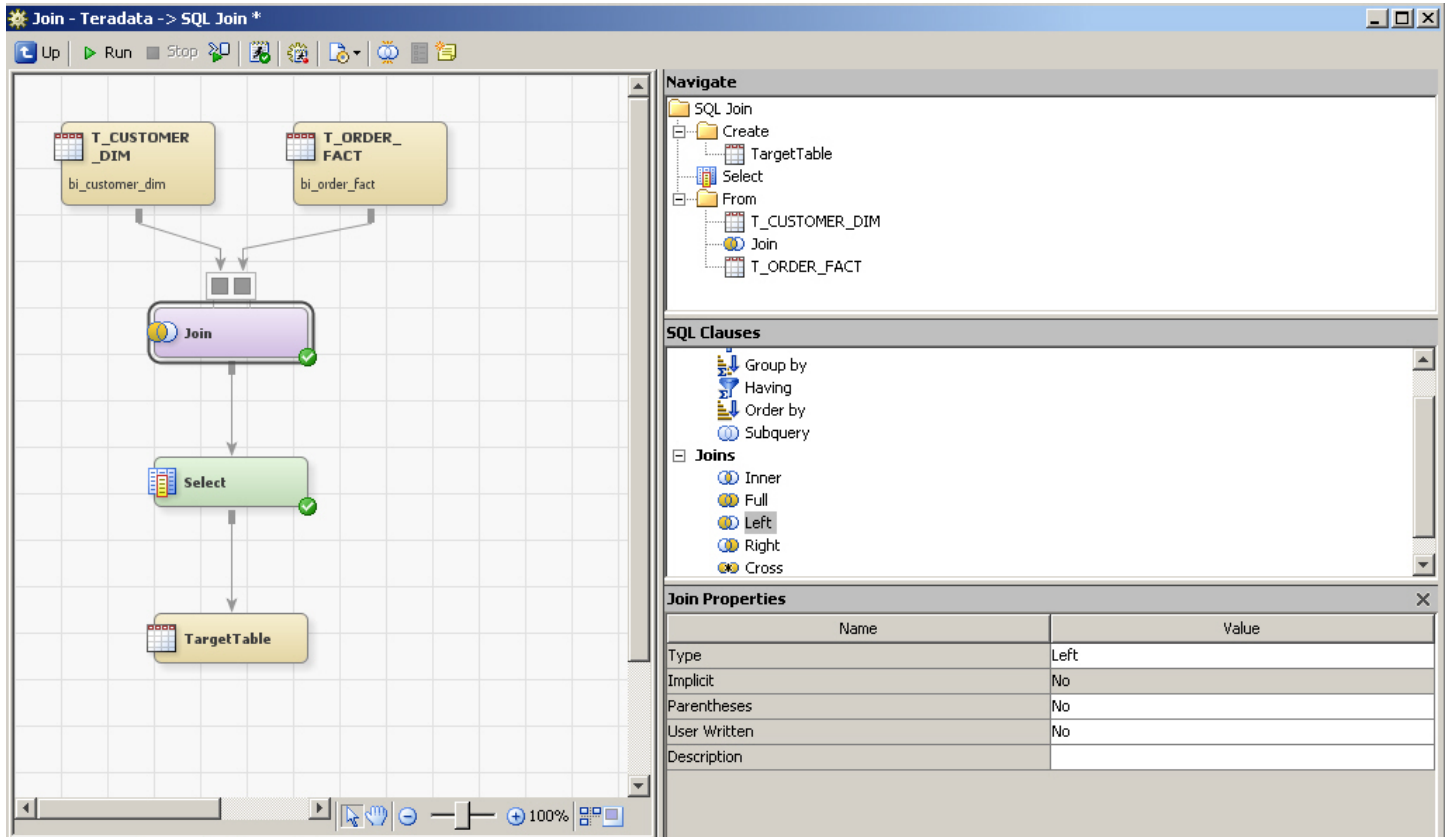


Apply the business rule within SAS Data Integration Studio.

## 4. Complex mapping

- Show how you would accomplish the equivalent of an outer join between two tables.

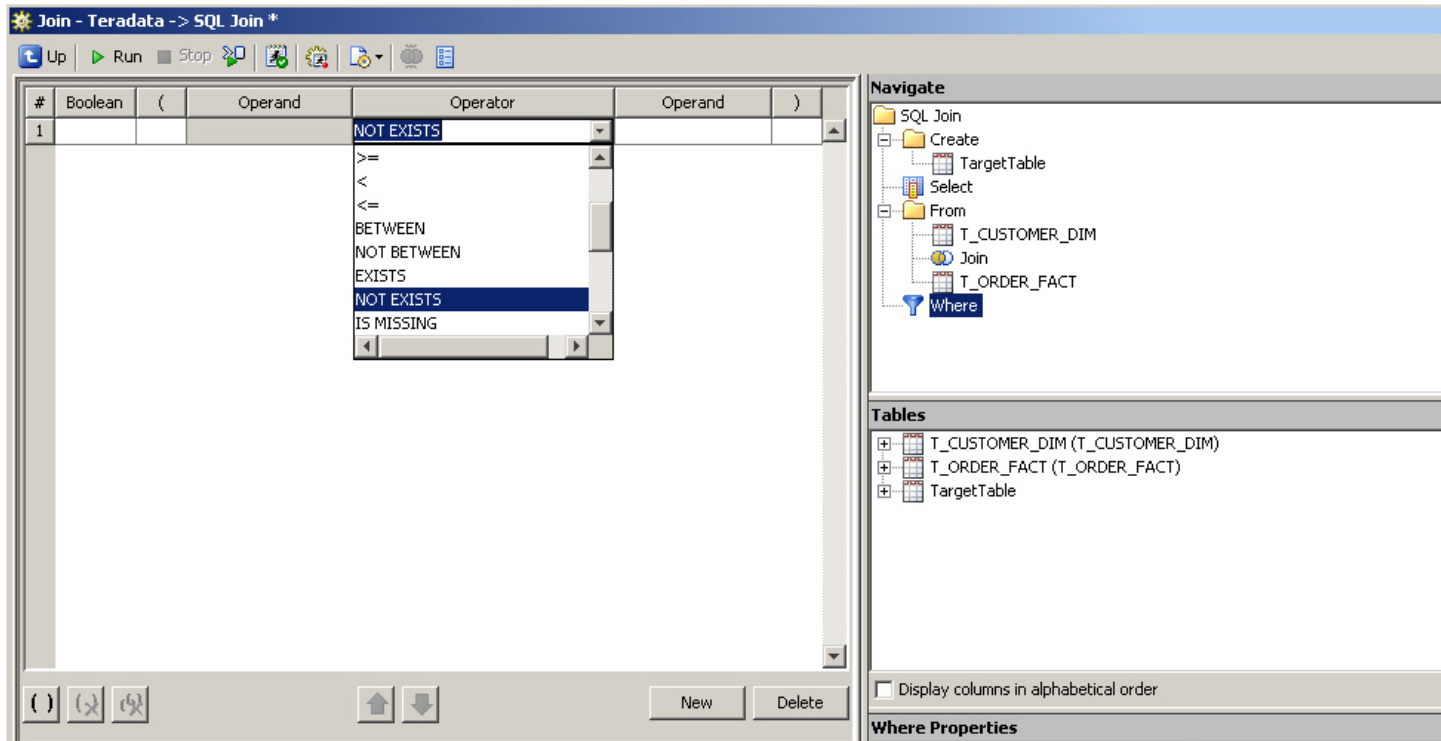
## Evaluating ETL Tools and Technologies



Various types of joins are supported and defined by selecting the join type within the SQL Designer in the SQL Properties window.

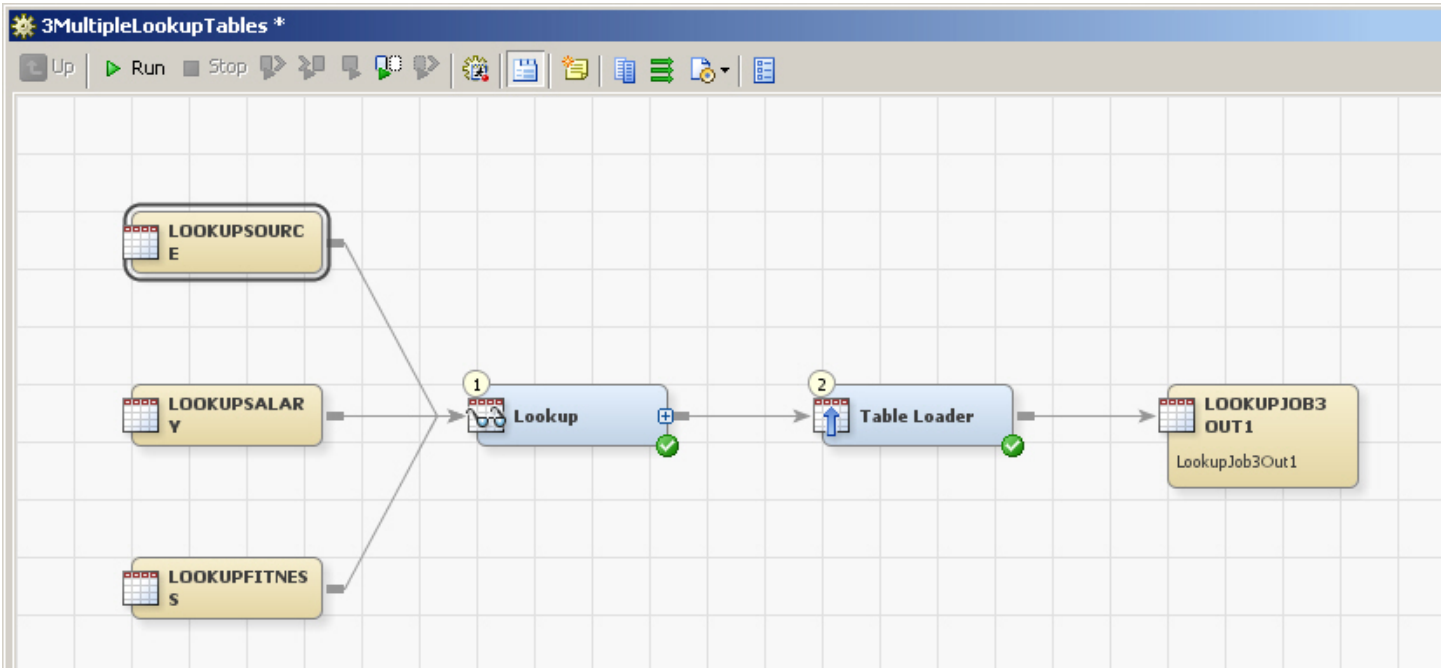
- Show how you would create an extract for data that is present in one table but is not present in another (equivalent to a “NOT EXISTS” clause in a SQL statement).

## Evaluating ETL Tools and Technologies



- Demonstrate a mapping that involves multiple sources, conditional lookups, and calculations to populate a target, as would be common for generating a fact table.

## Evaluating ETL Tools and Technologies

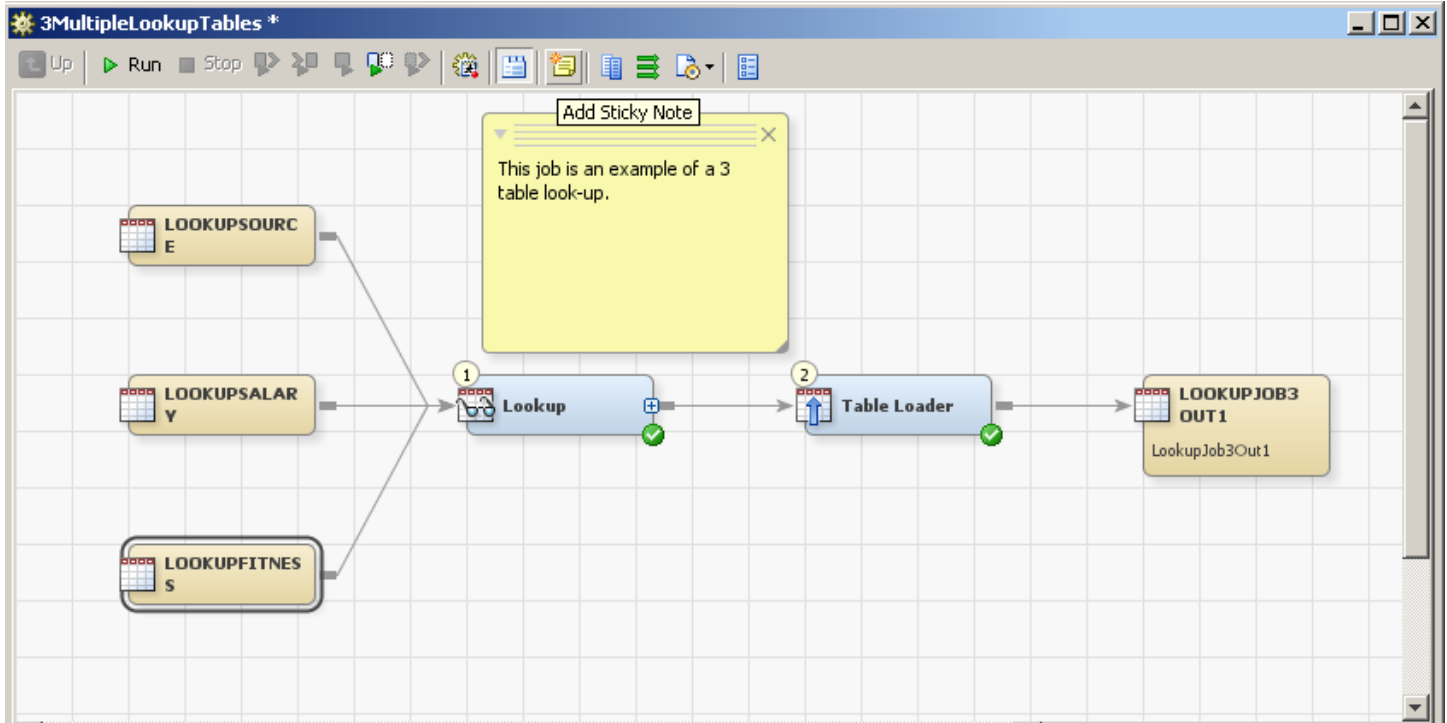


Here is an example of a process flow using the Look-up Transformation that completes a multiple sources look-up in a single pass. Conditional logic is defined to determine erroneous records and desired action based on the results.

### 5. Maintenance features

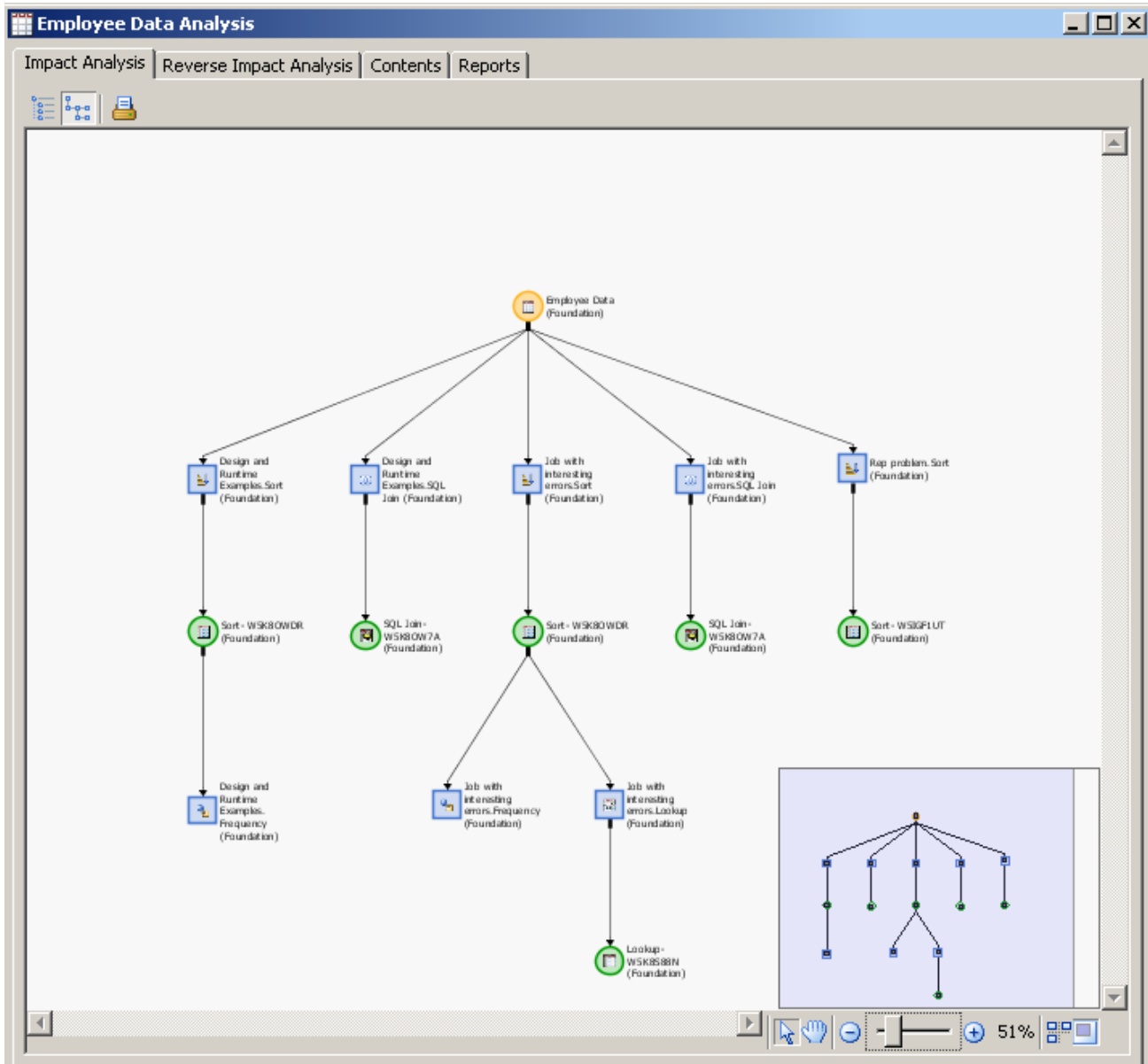
- Demonstrate the facilities available for documenting or annotating extracts.

# Evaluating ETL Tools and Technologies



- Demonstrate the process for tracing the impact of a source column / data change and updating the extract.

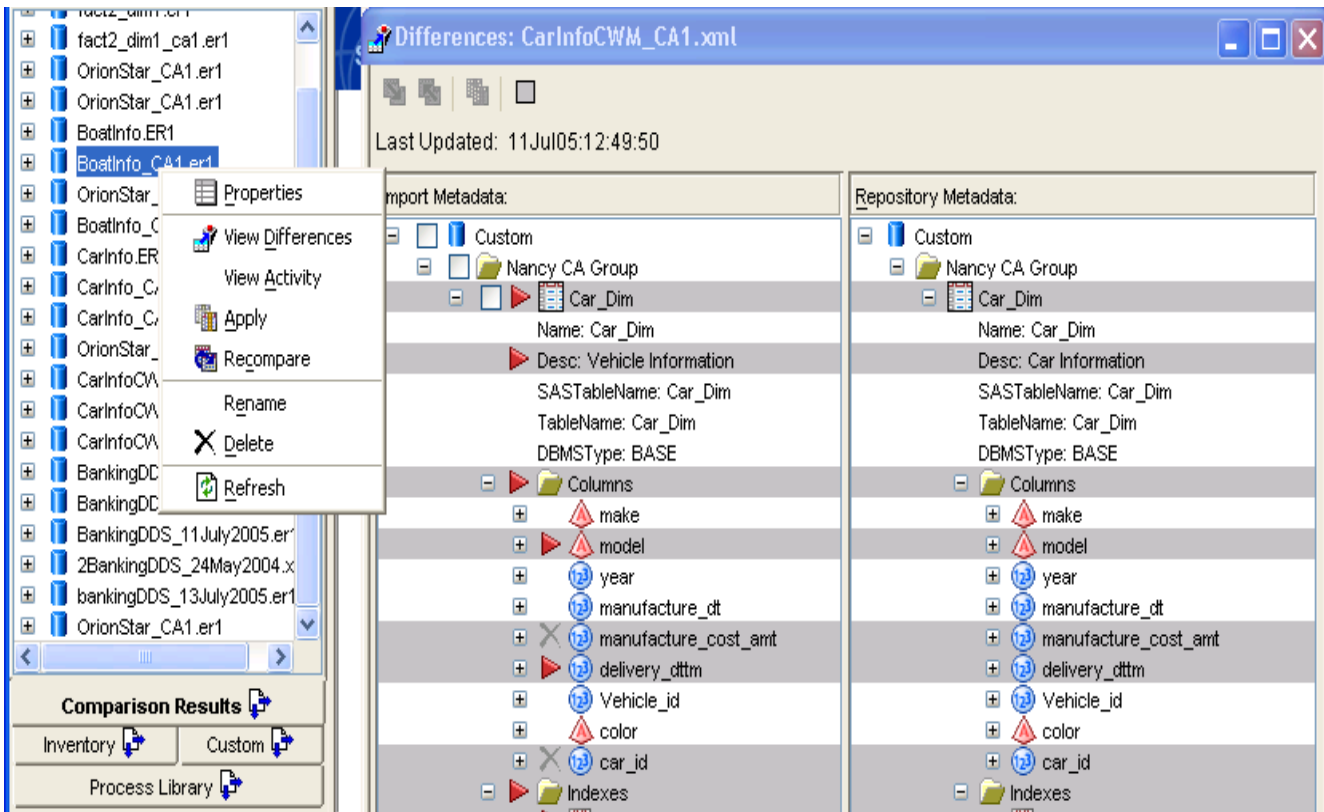
# Evaluating ETL Tools and Technologies



Impact Analysis and Reverse Impact Analysis (Data Lineage) is supported throughout the product including table and column level analysis. Impact Analysis can also reach outside the ETL environment to include business level metadata, OLAP Cubes and more.

# Evaluating ETL Tools and Technologies

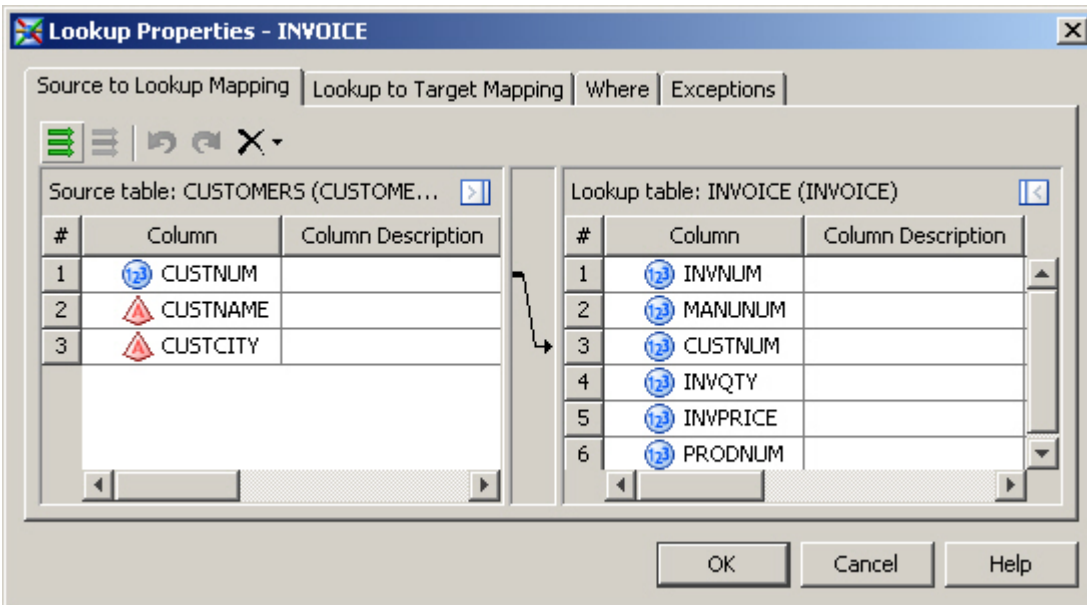
- Show how to detect differences between an extract in development and one that has been deployed in production.



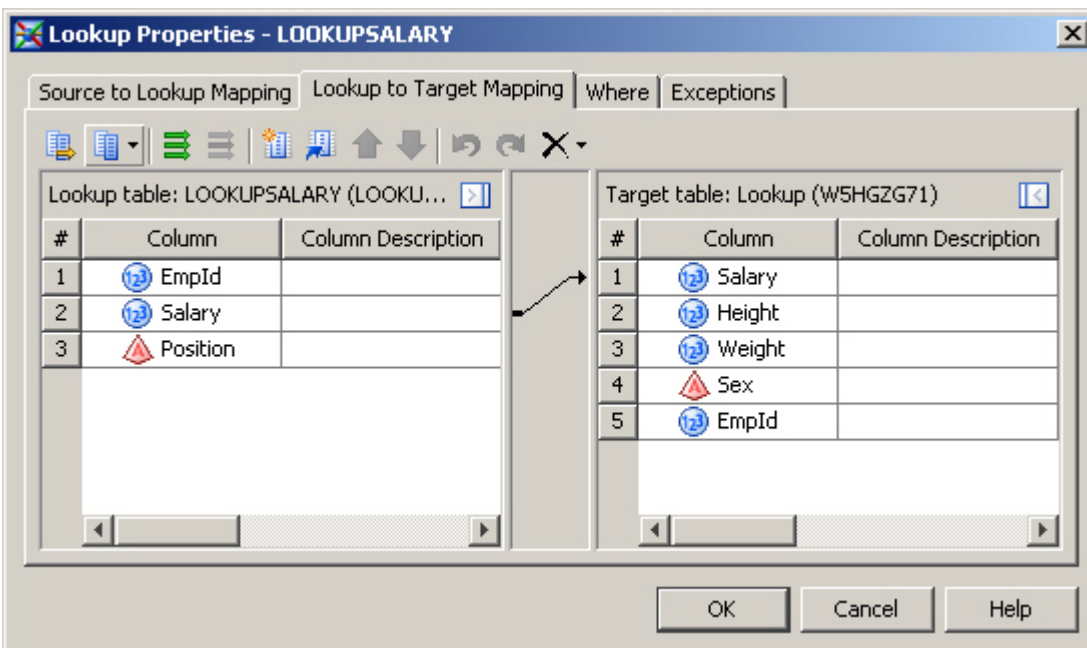
# Evaluating ETL Tools and Technologies

## 2. Lookups

- Show a simple lookup for something like a description for a code value.



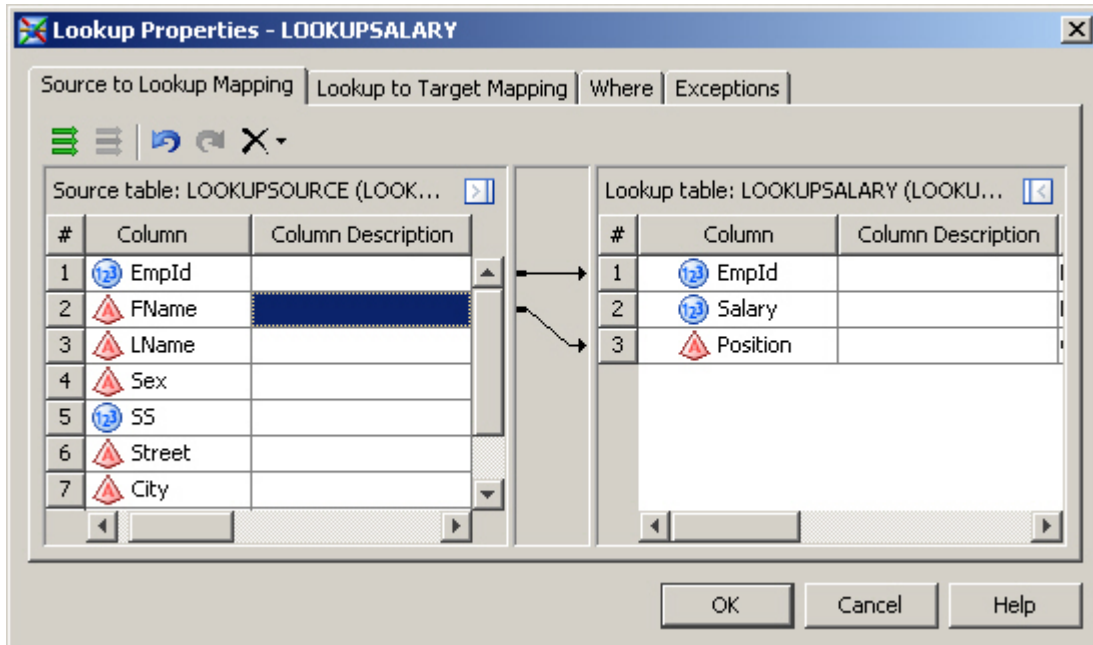
Simple look-up on single column key



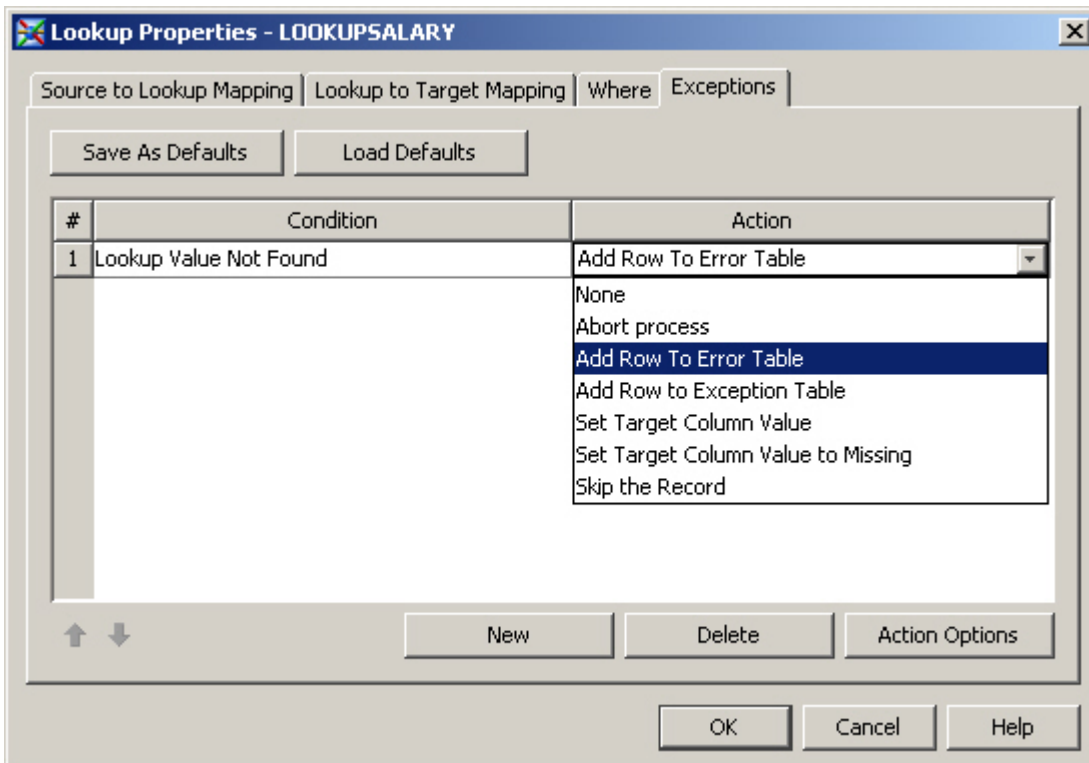
Extract look-up value from source and place into target table

- Show how to accomplish the same lookup, but assume you are retrieving two or more values from the lookup table rather than a single description.
- Demonstrate a lookup based on a multi-column input key.

## Evaluating ETL Tools and Technologies

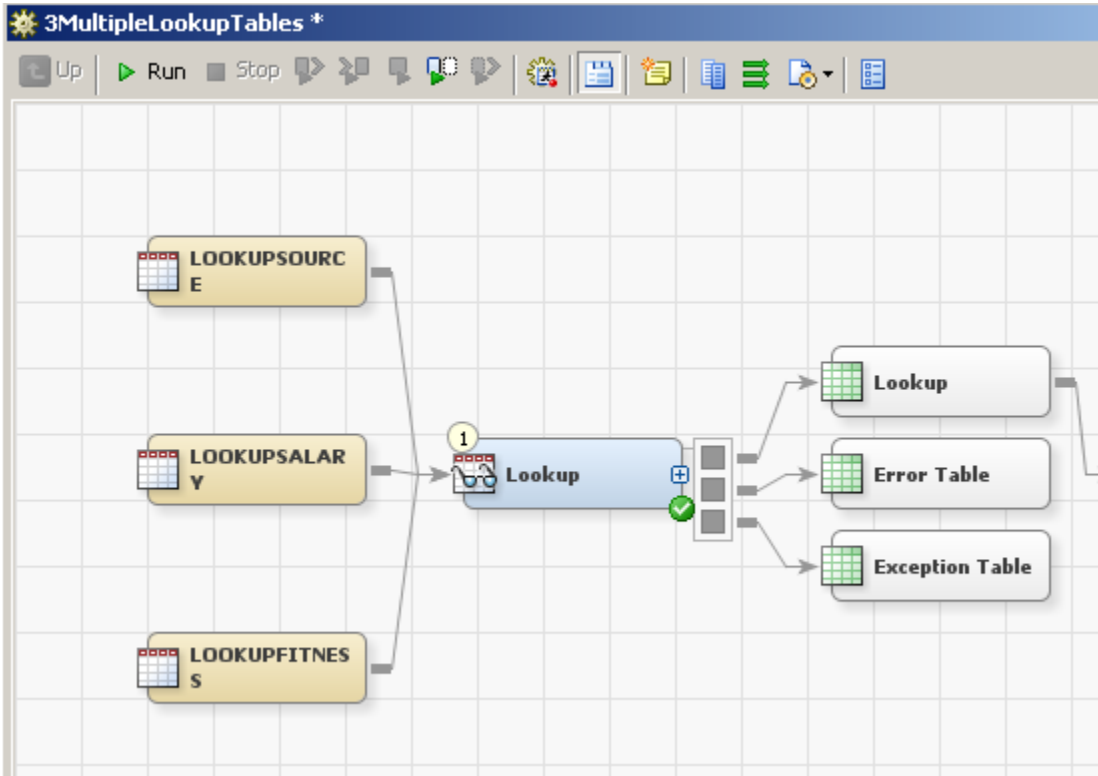


Look-up showing a multi-column look-up key



Customized error handling per look-up

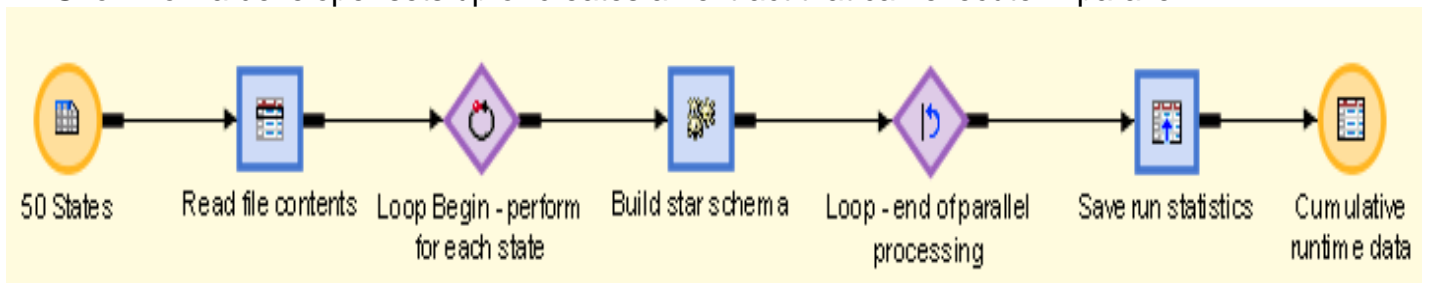
# Evaluating ETL Tools and Technologies



Creation of Error and Exception tables base on defined conditional criteria

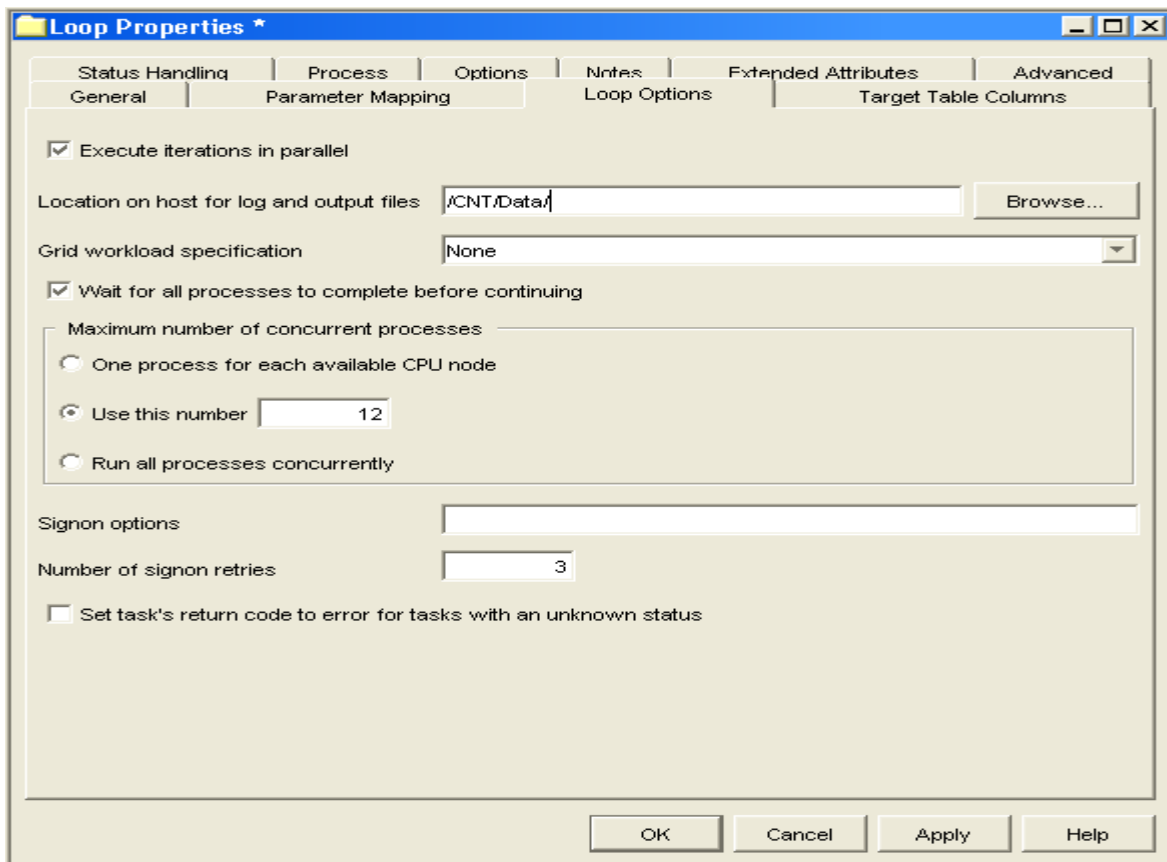
## 6. Performance

- Show how a developer sets up or creates an extract that can execute in parallel.



Parallel processing using partitioned data across multiple CPUs (SMP) or multiple servers (Grid)

## Evaluating ETL Tools and Technologies



Define sophisticated parallel processing options to control execution parameters

- Demonstrate or explain the features available for caching of source data or lookup values.

The Look-up Transformation does an in-memory cache of the look-up table.

This technique stores data in an expandable, in-memory table that has fast lookup capabilities for multiple source and target columns for multiple tables in a single pass of the data. It is similar to an array, where an index is used to store and retrieve data in the array. While an array uses numeric indices, a hash table can use character, numeric, or a combination of character and numeric indices. Unlike an array, the number of values stored in a hash table is not specified when creating the table. A hash table can grow to hold as many values as will fit into memory. When adding data to a hash table, an index, often called a key, is used to find a position in the table to copy the data. When retrieving data from a hash table, a key is given and a fast, nonlinear search occurs to determine if data with the key has been placed in the table. If so, the data stored with that key is returned.

- Describe or demonstrate the features available to deal with large volumes of source data.

SAS has over 30 years of experience in dealing with large volumes of data. SAS leverages many techniques to ensure optimal performance in a variety of computing environments with data at very large scales. Such capabilities include multi-threaded engines, native access to relational databases leveraging extract and load utilities, dynamic SMP/Grid enablement and server based

## Evaluating ETL Tools and Technologies

pooling. The techniques and others provide a highly scalable and integrated architecture for the enterprise from ETL to Business Intelligence.

- Explain how the product be configured to run extracts on more than one physical server, if this capability is available.

SAS is currently a member of the Global Grid Consortium and is actively expanding it's capabilities to better leverage multiple processing and grid environments. As shown above, SAS Data Integration Studio currently supports dynamic SMP and Grid enablement, distributing the work load to leverage all available computing resources. Data can also be access locally or remotely through native access engines and processed directly at its native environment.

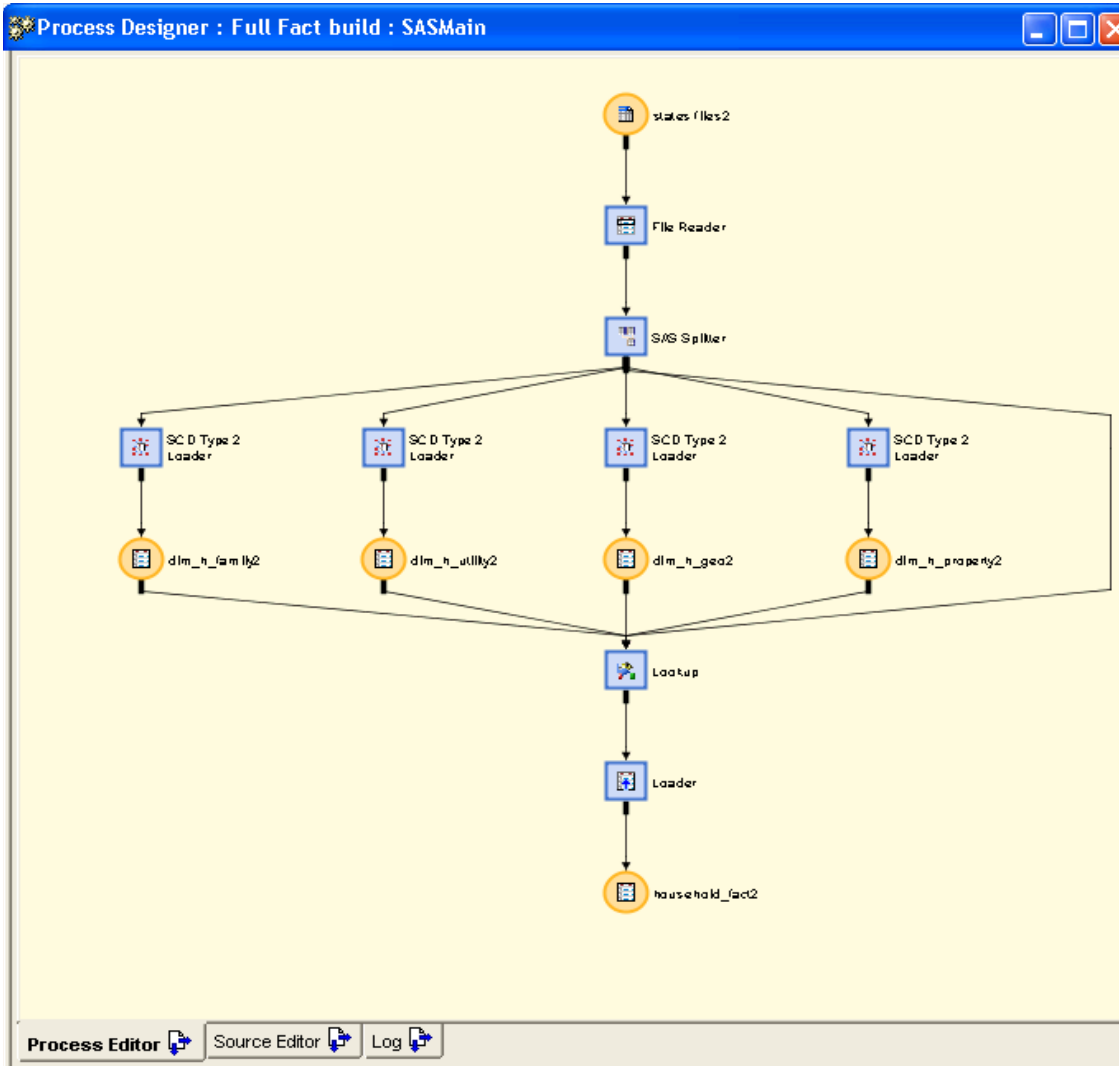
### **7. Scheduling and monitoring**

- What's the executable unit of work in the product?

Jobs are the basic unit of work. They can be run interactively or deployed multiple ways including batch scheduled, web service deployment and as embedded business logic.

- Show how you set up dependencies for tasks within an extract.

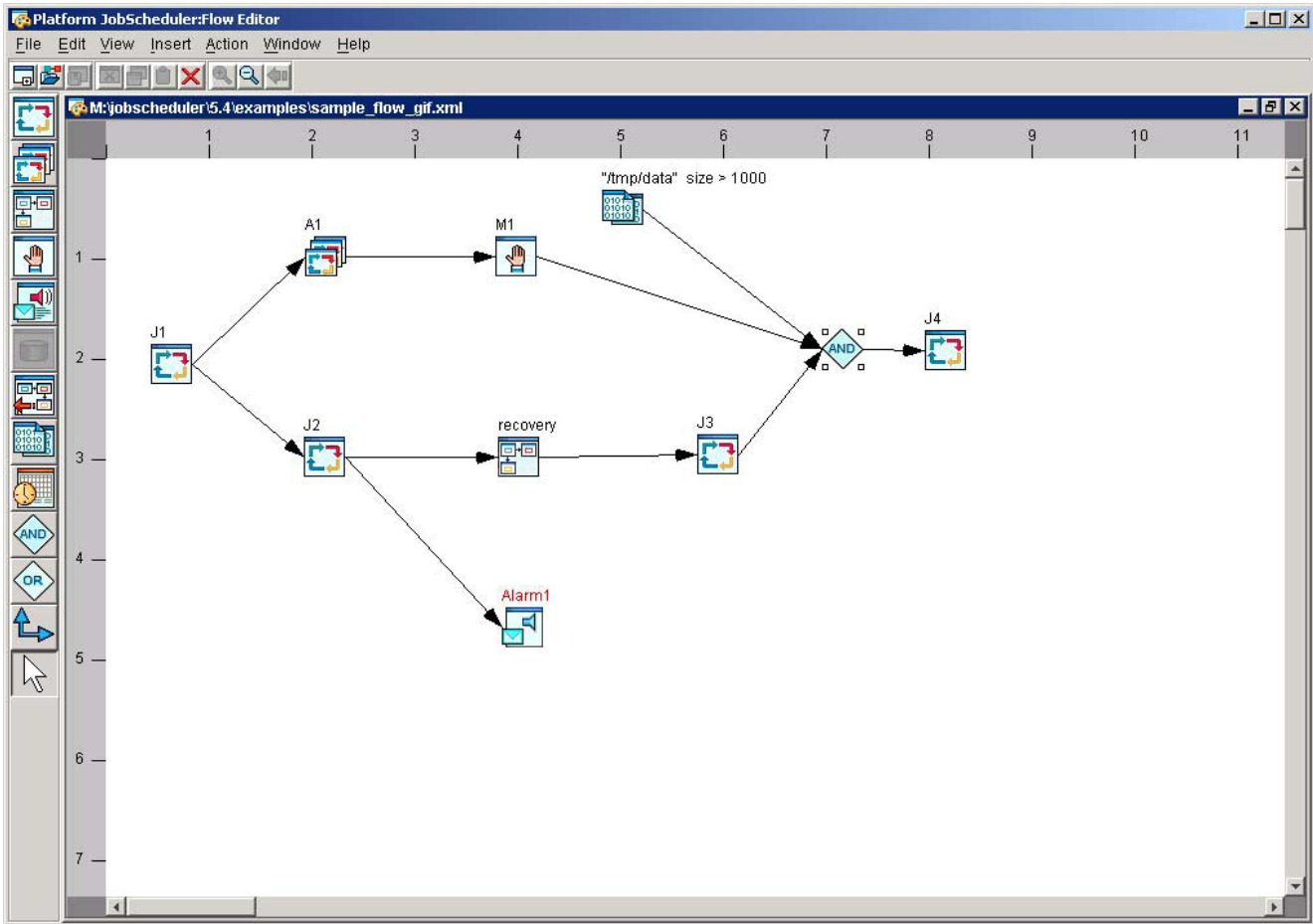
# Evaluating ETL Tools and Technologies



Steps within a job show the dependencies for each task

- Demonstrate how execution dependencies between jobs are set up (e.g. fact table doesn't load until dimension tables have loaded).

# Evaluating ETL Tools and Technologies



SAS partners with Platform Computing as a means to schedule jobs and define dependencies between jobs using the Job Flow Editor.

- Show how schedule-based initiation and event-based initiation of jobs works.

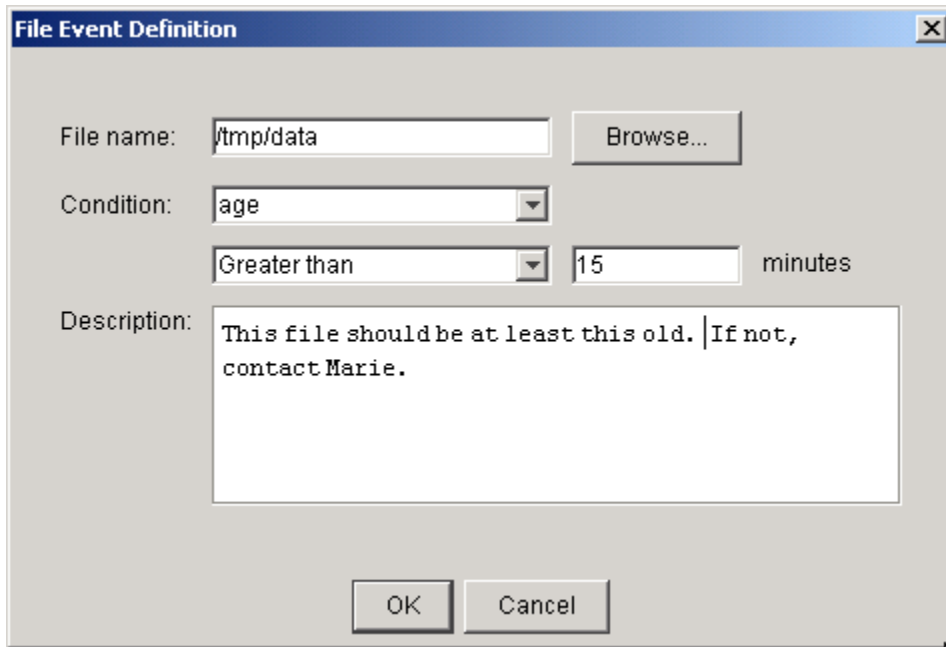
The Event Definition dialog box is used to specify the date and time for an event. The fields are as follows:

- Calendar name:** Businessdays@Sys (with a Refresh button)
- Hours:** 17 (with a note: (e.g \*, 1, 6, 1-7))
- Minutes:** 00 (with a note: (e.g \*, 10, 40, 20-30))
- Duration of event:** 60 minutes
- Description:** Must run between 5 and 6 p.m.

Buttons for OK and Cancel are located at the bottom of the dialog.

Time based dependencies

## Evaluating ETL Tools and Technologies

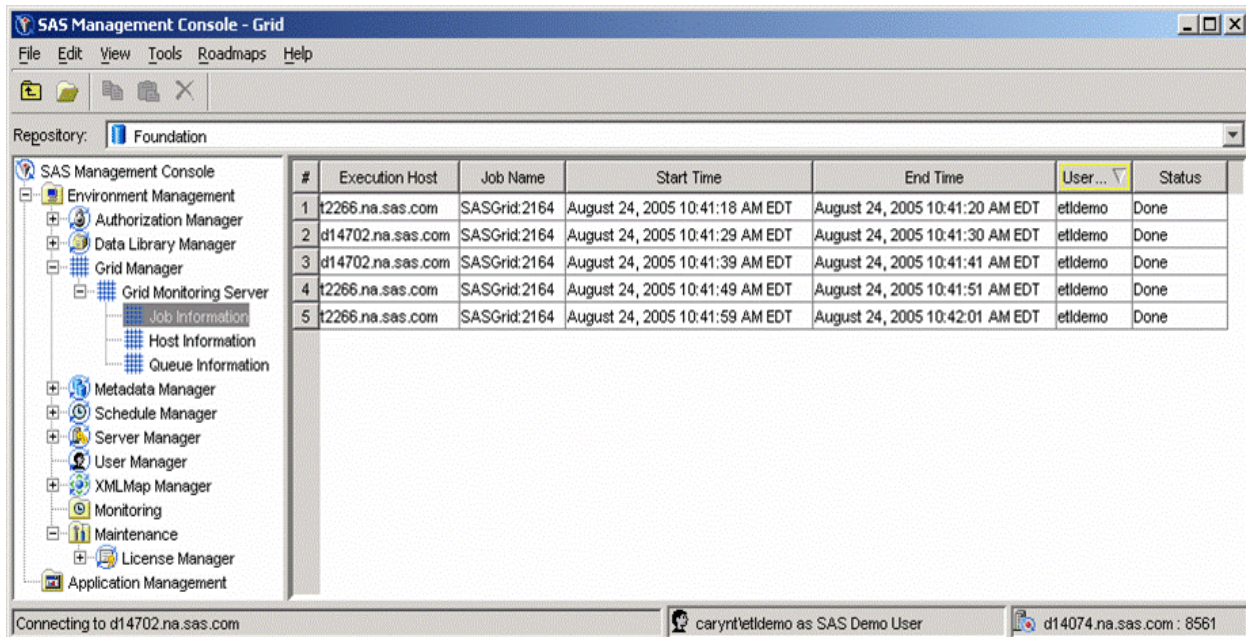


The dialog box is titled "File Event Definition" and contains the following fields and controls:

- File name:** A text input field containing "/tmp/data" and a "Browse..." button to its right.
- Condition:** A dropdown menu currently showing "age".
- Operator:** A dropdown menu currently showing "Greater than".
- Value:** A text input field containing "15" followed by the text "minutes".
- Description:** A text area containing the text: "This file should be at least this old. |If not, contact Marie."
- Buttons:** "OK" and "Cancel" buttons at the bottom.

### File base dependencies

- Show how execution can be monitored by a developer and by an administrator (outside the design environment).



The screenshot shows the SAS Management Console - Grid interface. The main window displays a table of job execution details. The table has the following columns: #, Execution Host, Job Name, Start Time, End Time, User, and Status. The data rows are as follows:

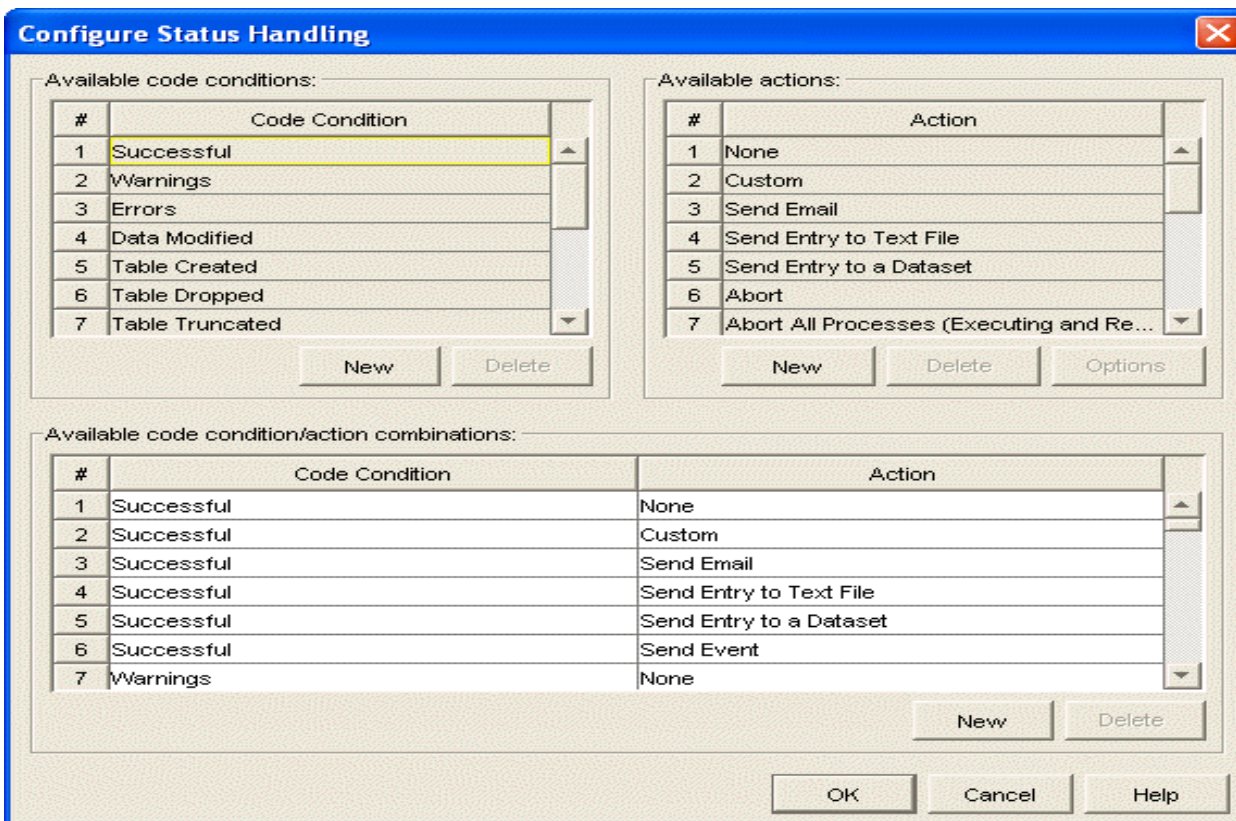
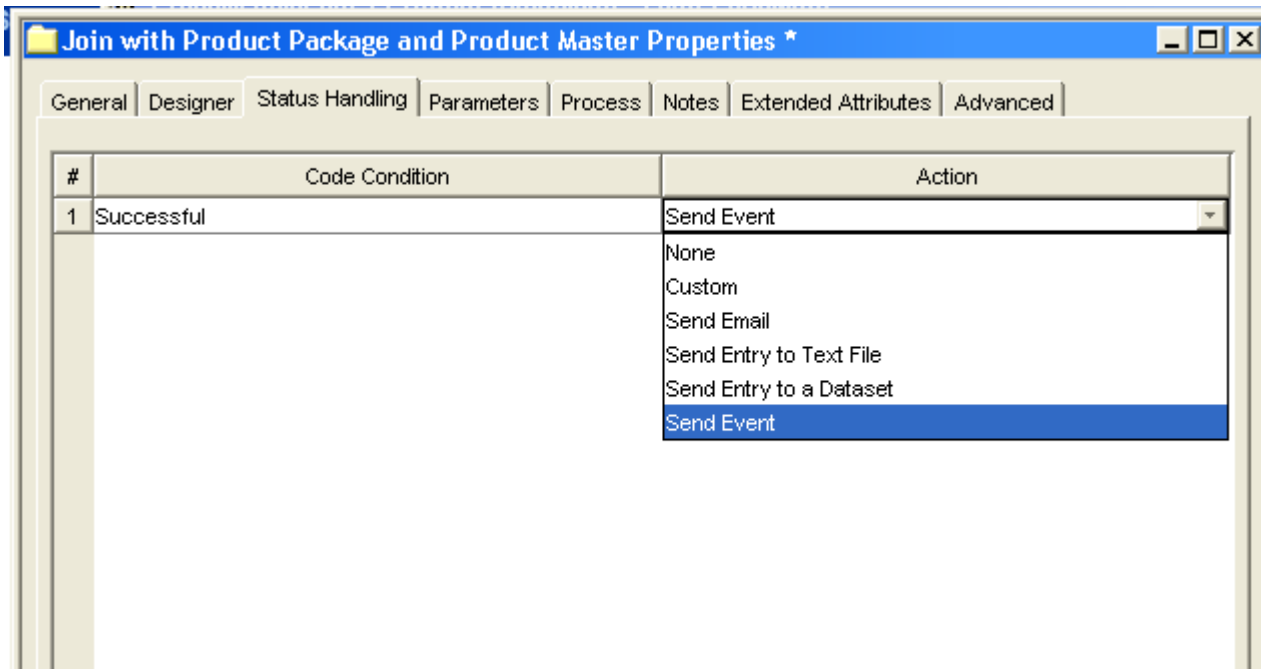
#	Execution Host	Job Name	Start Time	End Time	User	Status
1	t2266.na.sas.com	SASGrid:2164	August 24, 2005 10:41:18 AM EDT	August 24, 2005 10:41:20 AM EDT	etldemo	Done
2	d14702.na.sas.com	SASGrid:2164	August 24, 2005 10:41:29 AM EDT	August 24, 2005 10:41:30 AM EDT	etldemo	Done
3	d14702.na.sas.com	SASGrid:2164	August 24, 2005 10:41:39 AM EDT	August 24, 2005 10:41:41 AM EDT	etldemo	Done
4	t2266.na.sas.com	SASGrid:2164	August 24, 2005 10:41:49 AM EDT	August 24, 2005 10:41:51 AM EDT	etldemo	Done
5	t2266.na.sas.com	SASGrid:2164	August 24, 2005 10:41:59 AM EDT	August 24, 2005 10:42:01 AM EDT	etldemo	Done

The interface also includes a navigation tree on the left with categories like Environment Management, Grid Manager, and Metadata Manager. The status bar at the bottom shows the user is logged in as "carynt\etldemo as SAS Demo User" and is connected to "d14074.na.sas.com : 8561".

SAS Management Console provides job monitoring capabilities for any scheduled SAS job running on any execution host, including grid servers

## Evaluating ETL Tools and Technologies

- Explain the mechanisms available for monitoring execution and sending alerts.



Conditions can be defined at the job and transformation level to control the action taken based on specific events (process based, data based and user defined).