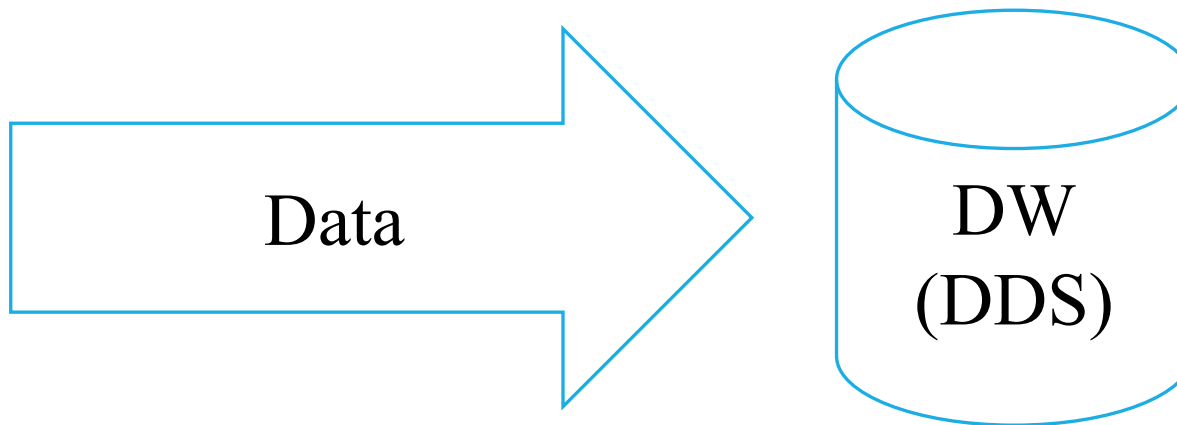




Data Delivery

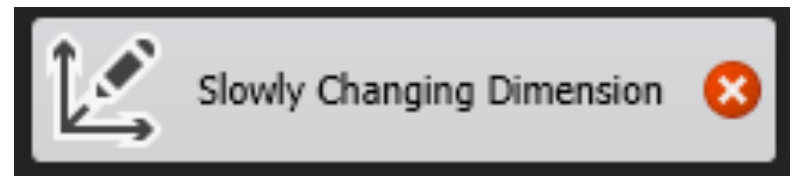
School of Information Studies
Syracuse University

Presentation Systems



Slowly Changing Dimension Manager

- The ETL system must determine how to handle a dimension attribute value that has changed from what is already in the warehouse.
 - Type 1: Overwrite
 - Type 2: New row
 - Type 3 = New column

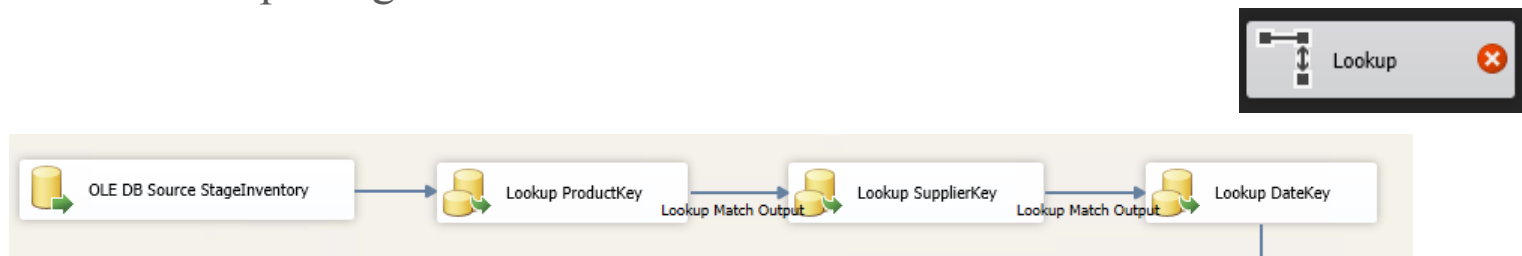


Surrogate Key Manager

- Surrogate keys are recommended for PKs of your dimension tables.
- In SQL Server, use **IDENTITY**.
- In other DBMSs a sequence with a database trigger can be used in place of identity.
- The ETL system can also manage them.

Surrogate Key Pipeline

- A system for replacing **operational natural keys** in the incoming **fact table record** with appropriate **dimension surrogate keys**.
- Approaches to handling **referential integrity** errors:
 - Throw away fact rows: bad idea
 - Write bad rows to an error table: most common
 - Insert placeholder row into the dimension: most complex
 - Fail the package and abort: draconian



Aggregate Builder

- Aggregates are specific data structures created to improve performance.
- Aggregates must be chosen carefully: Overaggregation is just as problematic as not enough.
- **Summary tables** and **subset dimensions** are generated from the base facts/dimensions.

