

Delivering a Relational Data Warehouse

Week 4 – Loading and Maintaining a Data Warehouse

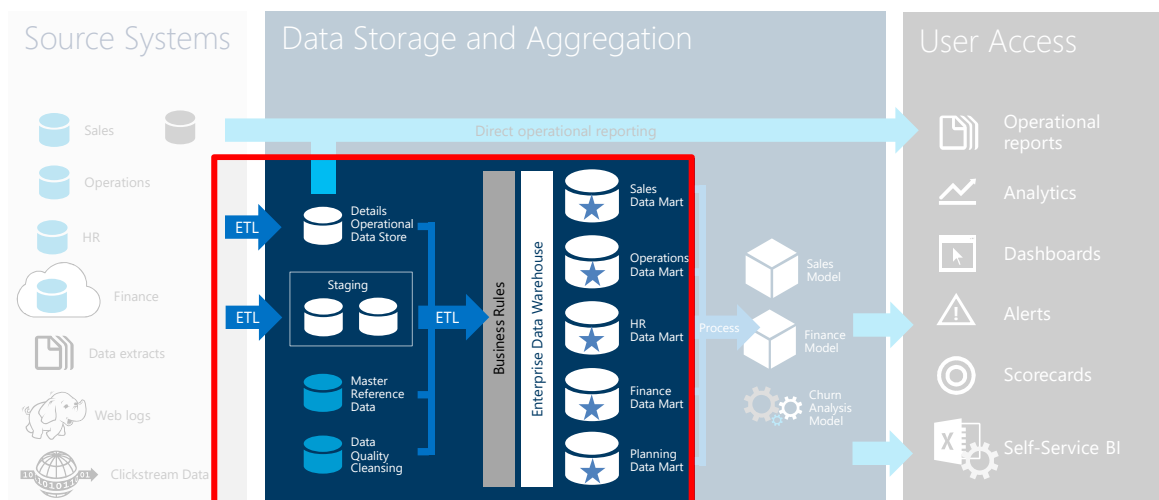
Module 10

Designing an Enterprise Information Management Strategy



Week Outline

4 | Loading and Maintaining a Data Warehouse



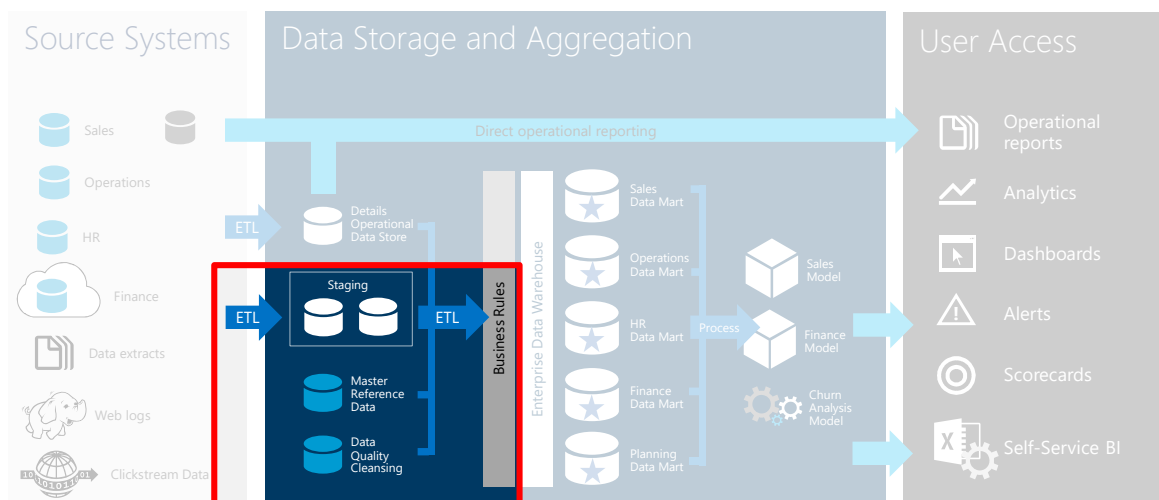
Module Outline

10 | Designing an Enterprise Information Management Strategy

Topic
▶ Enterprise Information Management
▶ Master Data Management
▶ Demo: Delivering MDM with Master Data Services
▶ Data Quality Management
▶ Demo: Delivering DQM with Data Quality Services

Module Outline

10 | Designing an Enterprise Information Management Strategy





©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Module Outline

10 | Designing an Enterprise Information Management Strategy

Topic
Enterprise Information Management
Master Data Management
Demo: Delivering MDM with Master Data Services
Data Quality Management
Demo: Delivering DQM with Data Quality Services

Enterprise Information Management

“Enterprise Information Management (EIM) is a best practice for creating, managing, sharing, and leveraging information in an enterprise, holistic manner that’s aligned with strategic, data-driven business objectives.”

– Philip Russom, TDWI

Enterprise Information Management

(Continued)

- EIM, more concisely, is about improving the “Four Cs” of data:
 - Completeness
 - Cleanliness
 - Consistency
 - Currency
- The overriding goal of EIM is to ensure that the investments made in data systems generate business value and mitigate risk

Enterprise Information Management

(Continued)

- EIM needs to be an enterprise-wide priority
- The responsibility of EIM efforts should fall at least as much on business as it does on IT—and preferably more
- It should strive to overcome traditional IT-related barriers to managing information at an enterprise level
- It should also provide a framework, tools, and vocabulary for delivering optimized value from data

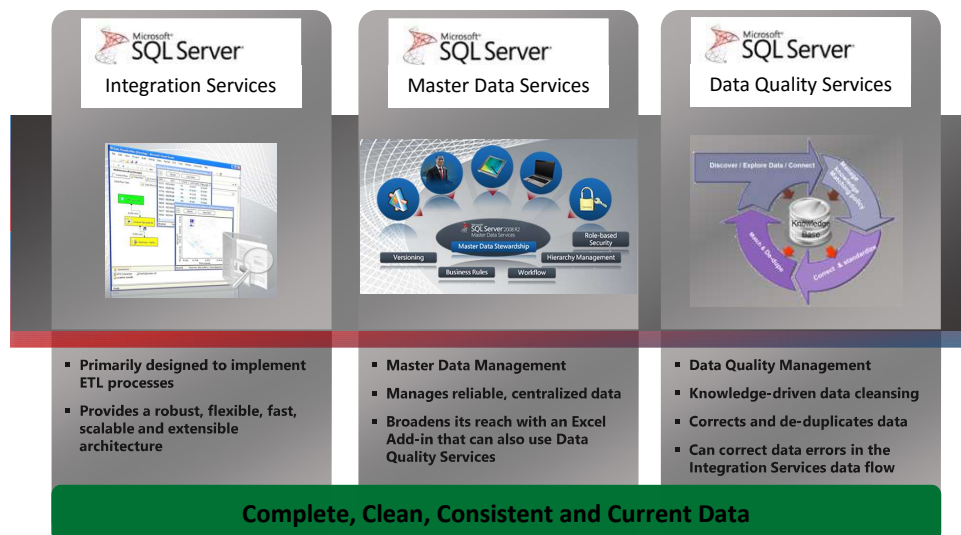
Enterprise Information Management

(Continued)

- EIM comprises several disciplines, including:
 - Master Data Management, and
 - Data Quality Management

Enterprise Information Management

SQL Server EIM Services



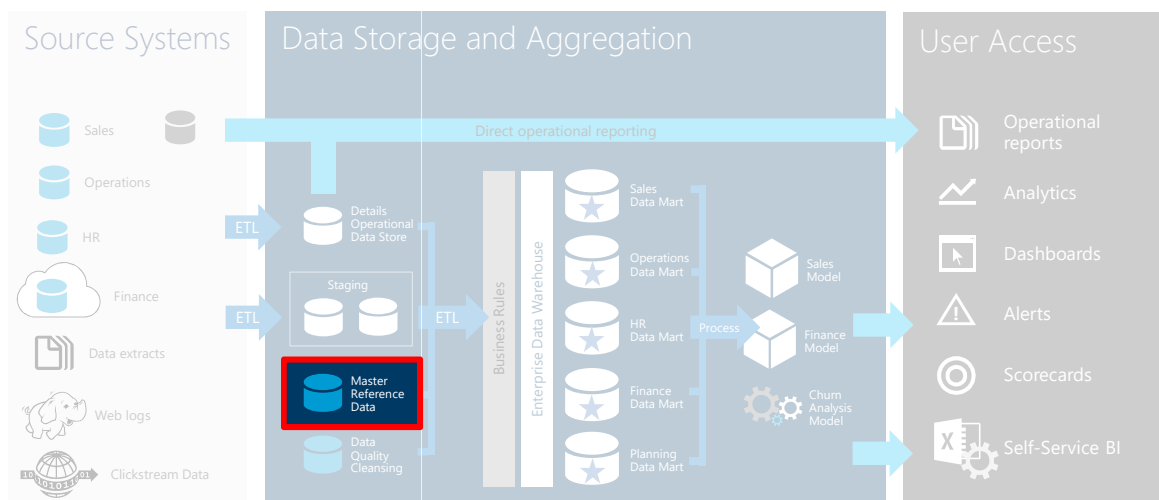
© 2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Module Outline

10 | Designing an Enterprise Information Management Strategy

Topic
Enterprise Information Management
Master Data Management
Demo: Delivering MDM with Master Data Services
Data Quality Management
Demo: Delivering DQM with Data Quality Services

Master Data Management



Master Data Management

"Master Data Management (MDM) comprises a set of processes, governance, policies, standards and tools that consistently defines and manages the master data (i.e. non-transactional data entities) of an organization (which may include reference data)."

– Wikipedia

Master Data Management

(Continued)

- MDM is about taking control of data, and making data work to the organization's advantage
 - A successful MDM solution delivers a master data hub that provides access to authoritative, standardized, and validated versions of data
- MDM can be responsible for deploying master data hubs that standardize data for key business entities across the enterprise
 - In the context of data warehousing, this refers to dimension records
 - Business entities typically refer to people, products, places and concepts

Master Data Management

(Continued)

- MDM can address several problems:
 - No authoritative source of data
 - ▶ Reports may not be trusted
 - ▶ Poor decision making
 - Systems with different versions of the same data
 - Data spread across multiple systems, lacking integrity and consistency
 - Mergers and acquisition that involve integrating data which contribute to different versions

Master Data Management

(Continued)

- MDM can address several problems:
 - Multiple stakeholders for the same record
 - Lack of formal dimension maintenance
 - An inability to enable, monitor and enforce data management processes

Master Data Management

SQL Server Master Data Services

- Master Data Services (MDS) is a service for master data management delivered with SQL Server
 - First released with SQL Server 2008 R2
 - Deployed as a WCF Data Service
 - The repository for metadata and data is a SQL Server database
 - Microsoft provides two user interfaces:
 - Master Data Manager web application
 - Excel add-in

Master Data Management

SQL Server Master Data Services ► Master Data Manager

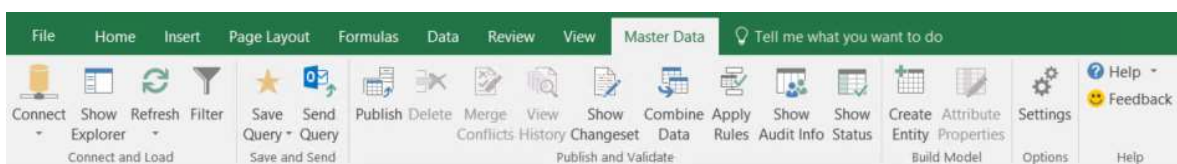
- Capabilities and features:
 - Work with master data:
 - Query, add, modify, delete, annotate master data
 - Manage hierarchy relationships
 - Administrative tasks:
 - Build models, create business rules and deploy model packages
 - Import data into MDS
 - Create subscription views to export data from MDS
 - Create versions of master data to manage consistent snapshots
 - Assign permissions to master data

Master Data Management

SQL Server Master Data Services ► Excel Add-in

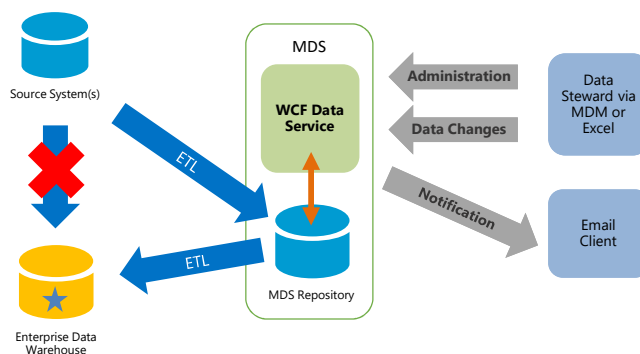
■ Capabilities and features:

- Connect to an MDS repository
- Create new entities and manage attribute properties
- Load filtered sets of master data into Excel tables
- Store and share queries to load master data
- Make changes to master data—and even work offline
- Publish and annotate changes



Master Data Management

SQL Server Master Data Services ► Sample Implementation



Master Data Management

SQL Server Master Data Services ► Development Methodology

1. Create a model
2. Create entities and attributes
3. Define business rules
4. Define hierarchies
5. Import data to initialize/populate entity memberships
6. Create subscription views to export master data
7. Configure permission requirements
8. Manage model versioning



©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Module Outline

10 | Designing an Enterprise Information Management Strategy

Topic
Enterprise Information Management
Master Data Management
Demo: Delivering MDM with Master Data Services
Data Quality Management
Demo: Delivering DQM with Data Quality Services

Demo

Delivering MDM with Master Data Services

Demo objectives:

1. Develop an MDS solution
2. Manage an MDS solution



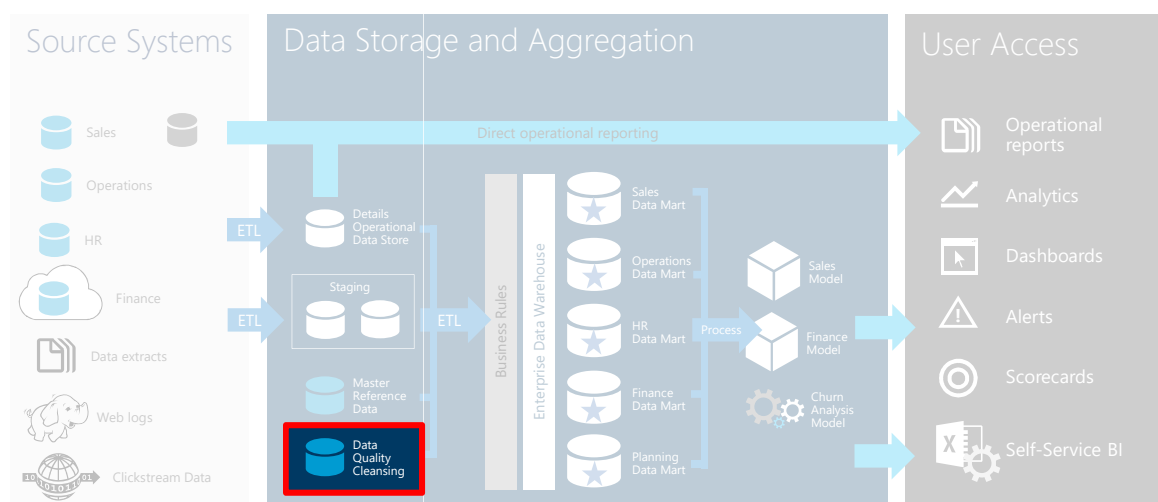
©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Module Outline

10 | Designing an Enterprise Information Management Strategy

Topic
Enterprise Information Management
Master Data Management
Demo: Delivering MDM with Master Data Services
Data Quality Management
Demo: Delivering DQM with Data Quality Services

Data Quality Management



Data Quality Management

Data Quality Assurance

"Data Quality Assurance (DQA) is the process of profiling the data to discover inconsistencies, and other anomalies in the data and performing data cleansing activities (e.g. removing outliers, missing data interpolation) to improve the data quality.

These activities can be undertaken as part of data warehousing or as part of the database administration of an existing piece of applications software."

– Wikipedia

Data Quality Management

- Essentially, Garbage In ► Garbage Out
 - DQA processes detect and address garbage data
- Common processes include:
 - Data profiling
 - Data correction
 - Data standardization
 - Geocoding, for name and address data
 - Matching

Data Quality Management

SQL Server Data Quality Services

- Data Quality Services (DQS) is a knowledge-driven data cleansing service delivered with SQL Server
 - First released with SQL Server 2012
 - It enables the cleansing, matching, standardization of data
 - Data cleansing can be performed in a Data Quality Project
 - Results are output to a SQL Server table, Excel workbook or CSV file
 - Data cleansing can be integrated with:
 - Integration Services: DQS Cleansing transform
 - Master Data Services: Matching in the Excel MDS Add-in

Data Quality Management

SQL Server Data Quality Services (Continued)

- DQS consists of a Data Quality Server and the Data Quality Client
- The Data Quality Server is a SQL Server database instance feature that consists of three catalogs with data-quality functionality and storage

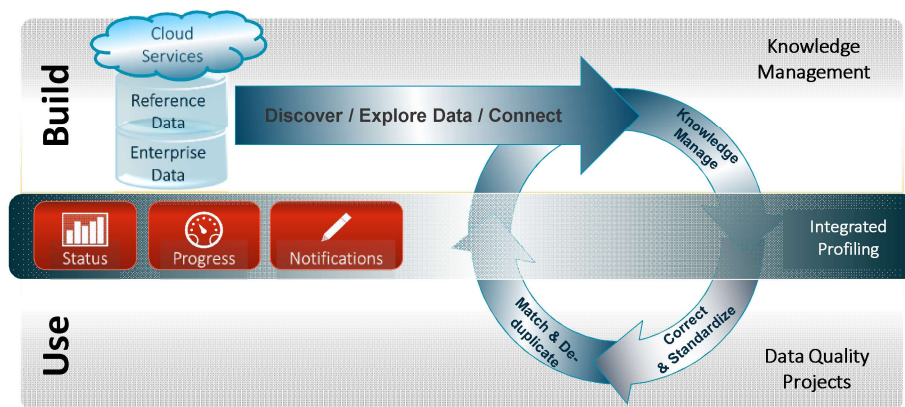
Data Quality Management

SQL Server Data Quality Services (Continued)

- The Data Quality Client is a standalone Windows application that allows data stewards, data experts, and IT Professionals to perform data quality operations, and supports:
 - Creating and maintaining knowledge bases
 - Creating and maintaining Data Quality Projects
 - Monitoring data quality activities
 - Managing system settings

Data Quality Management

SQL Server Data Quality Services ► Data Quality Process



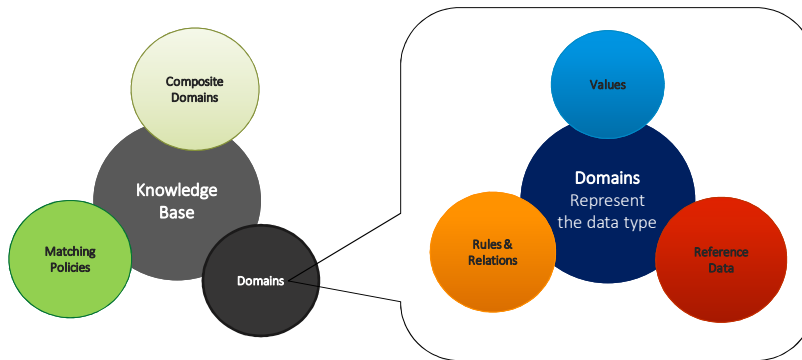
Data Quality Management

SQL Server Data Quality Services ► Knowledge Bases

- Rationale: To cleanse data you need knowledge about it!
- A knowledge base is a data repository of knowledge that enables professionals to understand their data and maintain its integrity
 - Knowledge in a knowledge base is maintained in domains, each of which is specific to a data field
 - Domains capture the semantics of the data
 - Domains can use online reference data
 - Online DataMarket Reference Data Service
 - Direct Online 3rd Party Reference Data Services

Data Quality Management

SQL Server Data Quality Services ► Knowledge Bases



Data Quality Management

SQL Server Data Quality Services ► Development Methodology

1. Create a knowledge base
2. Create and configure the knowledge base domains
3. Use the knowledge base in Data Quality Projects, and/or integrate in an Integration Services package data flow
4. Configure permission requirements
5. Continue to maintain and improve the knowledge in the knowledge base



©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Module Outline

10 | Designing an Enterprise Information Management Strategy

Topic
Enterprise Information Management
Master Data Management
Demo: Delivering MDM with Master Data Services
Data Quality Management
Demo: Delivering DQM with Data Quality Services

Demo

Delivering DQM with Data Quality Services

Demo objectives:

1. Create a knowledge base
2. Create a Data Quality Project



©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.