

# ORACLE DATA PROFILING AND QUALITY

## KEY FEATURES AND BENEFITS

DATA PROFILING AND DATA CLEANSING IN THE ORACLE DATABASE

### FEATURES

- Data Profiling
- Data Rules
- Data Auditing
- Automated Data Cleansing

### BENEFITS

- Know your data, discover problem sources early
- Define data rules once, use them everywhere
- Profiling and auditing statistics let you quantify ROI on DQ
- Discover and correct complex data problems without writing any code
- Flexible for use with any ETL solution, protecting data integration investments
- Best integration with Oracle Warehouse Builder
- Leverages database resources and skills to lower TCO

*Use the Oracle Data Profiling and Quality option to discover and define data rules, monitor data rule compliance, and cleanse dirty data. You can profile and cleanse data in any data store that the Oracle Database can access- including tables and views in Oracle and non-Oracle databases, files, and applications such as Siebel, eBusiness Suite, and SAP R/3. This option is optimized for use with Oracle Database Enterprise Edition.*

### Oracle Data Profiling and Quality

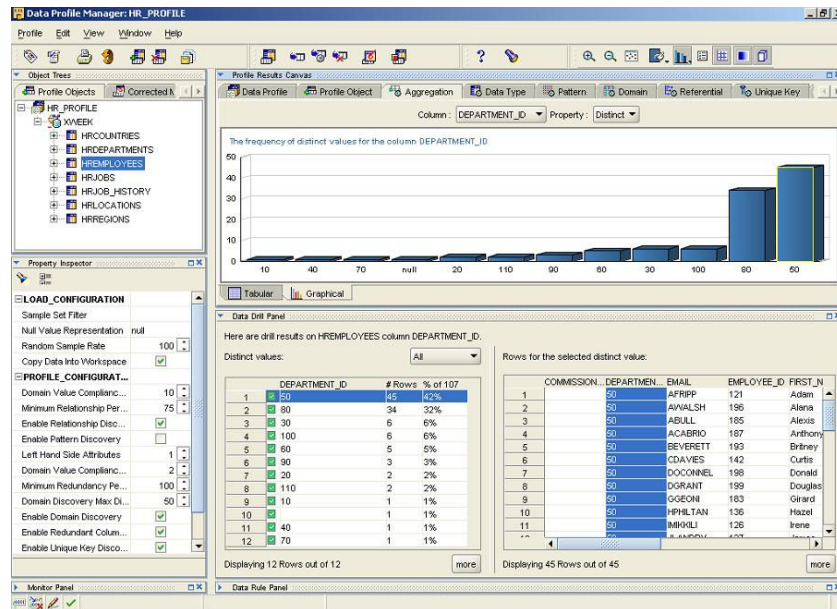
You can use Oracle Data Profiling and Quality to address the whole data quality lifecycle.

- Data profiling, for initial data quality assessment and relationship discovery
- Data rules, for active enforcement of data quality
- Data auditors, to monitor data quality based on data rules
- Data cleansing and correction, for automatic repair of defective data as identified by data rules

Oracle recommends that you profile and cleanse data as a precursor to attempting complex and time-consuming ETL transformations. Therefore, the Oracle Data Profiling and Quality option is available in the data warehousing and ETL tool embedded in the Oracle Database, which is Warehouse Builder.

### Data Profiling: Data Quality Discovery

Oracle Warehouse Builder (OWB) provides the user interface for profiling and exploring your data. The Data Profile Manager lets you visualize your data in the aggregate, view statistical quality measures like Six Sigma scores, check discovered data domains and relationships against your documented metadata, and drill down into the underlying details to view compliant and non-compliant data.



### Data Profiling UI in Warehouse Builder

Profiling metadata and statistics are captured in the OWB metadata repository to drive your data rules, ongoing data quality auditing, and automated data cleansing processes. You can also define data rules specific to your business or your problem domain, and test compliance with those rules during the data profiling process.

### Data Rules: Active Data Quality

OWB data rules turn patterns and relationships discovered in data profiling into standards that can be tested, monitored and enforced across all of your data sources and targets. You can also take your own business rules and represent them as data rules. Data rules are represented with the other metadata for your database in the OWB repository.

Data rules let you build an additional level of consistency checking into your data integration solution, beyond what can be enforced by database-level measures such as constraints. Data rules can span relationships among multiple sources or targets, and can include arbitrarily complex logic. Rules can be enforced inline during data loading and cleansing processes, or tested in out-of-band auditing processes.

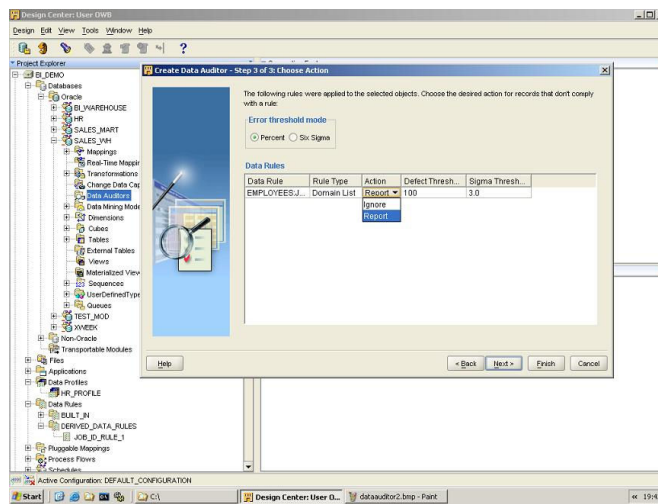
If you use OWB for your primary data integration tool, then you can use data rules directly in your ETL mappings. Apply rules case-by-case as needed, or bind data rules to an object, so you can enforce them wherever the object is used in ETL. Filter non-compliant data for special handling such as cleansing or enrichment at any point in your data integration process.

Representing data rules in the OWB repository ensures that they can be enforced consistently across your entire data integration design. If you update a rule, those changes can be propagated throughout all applications of the rule automatically—just regenerate your code and deploy the updated version.

### Data Auditors: Data Quality Monitoring

Data auditors test compliance of source or target data against your data rules, as part of data loading processes, at times you schedule, or on demand. They let you spot data quality issues with your sources and targets as they emerge, and monitor data quality metrics to track ROI on your data quality efforts.

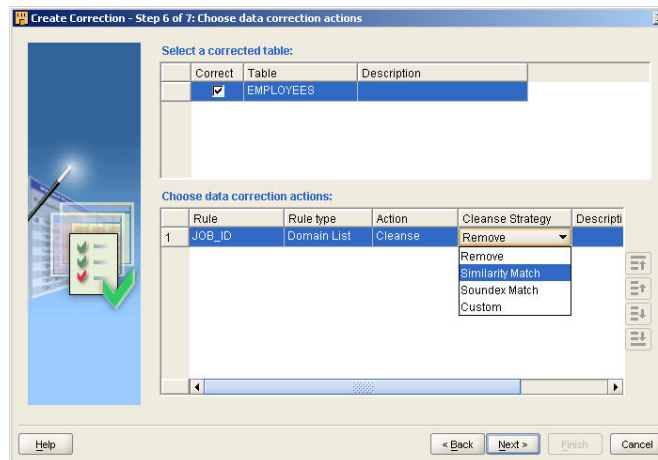
If the quality goes below a specified threshold (measured as a Six Sigma score or % defects value), the data auditor will warn you and you can take appropriate action. Data auditor results are stored in the OWB repository, and can be used as the basis of reports to track data quality over time.



Creating a Data Auditor

### Data Cleansing: Automatically Create Data Correction Processes

To cleanse data with minimum developer effort, OWB automates the creation of data correction mappings.



Specifying data cleansing strategies for data

Pick your source data, a given rule, and select from among advanced out-of-the-box

**RELATED PRODUCTS AND SERVICES**

Discover data patterns and relationships, assess data quality, define data rules, monitor data compliance, and cleanse bad data, all in one solution that works alongside any ETL tool on the market and leverages your existing resources and skills.

**RELATED PRODUCTS**

The following data quality and data integration products complement Oracle Data Profiling and Quality:

- Oracle Database Enterprise Edition
- Oracle Data Integrator Enterprise Edition
- Oracle Warehouse Builder
- Application Adapters for Oracle Warehouse Builder
- Oracle Data Quality for Data Integrator
- Oracle Data Profiling
- Oracle Master Data Management
- Oracle Gateways

correction algorithms, and OWB generates mappings that cleanse the data and write it to a target location you choose. You can also implement custom cleansing strategies, and OWB will still provide the logic to extract dirty data from the source and deliver cleansed results to the target. You then have the option of creating further mappings to update the source with cleansed results if your application requires it.

While these mappings use OWB ETL features to move data from the dirty source to the cleansed target, they do not require that you replace your other ETL processes

**Best for Oracle: In-Database Data Quality**

All of the above data quality features run within the Oracle Database and are available for all data stores that Database can access- including flat files, non-Oracle databases, and applications. Performing data quality in Oracle Database ensures high performance and leverages your existing database hardware investment and DBA and developer skill sets.

**Contact Us**

For more information about the Oracle Data Profiling and Quality option, please visit [oracle.com](http://oracle.com) or call +1.800.ORACLE1 to speak to an Oracle representative.



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2009, Oracle and/or its affiliates. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners. 0109