ORACLE

ORACLE ENTERPRISE MANAGER 13c

Oracle Data Masking and Subsetting Pack

Growing security threats and ever-expanding privacy regulations have made limiting the exposure of sensitive data an imperative for organizations. Copying production data for non-production purposes such as development, testing, and data analytics can cause sensitive data to proliferate, expanding the security and compliance boundary and increasing the likelihood of data breaches. Oracle Data Masking and Subsetting provides a flexible solution that discovers and masks sensitive data and can reduce the size of large data sets, enabling safe and resource-efficient data sharing across the extended enterprise.

Key Business Benefits

- Reduces sensitive data exposure in non-production environments.
- Improves compliance with data privacy laws and standards.
- Facilitates sharing of production data for test, development, data analytics, and third-party business partners.
- Minimizes storage costs by subsetting data.

INTRODUCTION TO DATA MASKING AND SUBSETTING

Non-production environments such as test and development systems are potential targets for cyber attacks because they generally contain copies of production data. These systems are typically not as well protected or monitored as production systems, increasing risks to your sensitive data. Limiting the amount of data copied and deidentifying it using data masking is recognized as a best practice for supporting non-production purposes.

Figure 1. Overview of Data Masking

Oracle Data Masking and Subsetting improves security by reducing the exposure of sensitive data in non-production environments. Data Masking and Subsetting extracts entire copies or subsets of application data from a database and masks sensitive data so it may be safely shared to support activities such as software QA and business analysis. Masking can help reduce costs by keeping masked non-production databases out of the scope of compliance audits.

Masking and subsetting data can also help maintain privacy and minimize network and storage requirements when sharing data with third parties and partners.

SENSITIVE DATA DISCOVERY AND MODELING

Finding sensitive data in complex applications is a non-trivial task. Application Data Modeling, a Data Masking and Subsetting feature, automates the discovery of columns likely to contain sensitive data and their corresponding parent-child relationships. The discovery process uses built-in extensible patterns such as credit card numbers and national identifiers to check metadata and column data to identify sensitive columns. The resulting Application Data Model provides a complete set of sensitive columns and referential relationships, ensuring that the masking and subsetting process maintains the application integrity of the data.

Key Features

- Automated discovery of sensitive columns and parent-child relationships
- Comprehensive and extensible built-in masking formats
- Creation and reuse of custom templates for applications
- Integrated data subsetting for data reduction
- Masking and subsetting in-database or during extraction
- Supports databases both on-premises and in the cloud
- Facilitates high-performance, repeatable masking and data minimization processes

DATA MASKING

Data Masking and Subsetting provides a comprehensive and extensible library of masking formats that define the transformations and logic used to mask data. Predefined masking formats streamline masking of sensitive data such as credit card numbers, national identifiers, and other personally identifiable information (PII).

Data Masking and Subsetting also allows you to quickly create new masking formats to meet your specific requirements. You can define simple masking formats using options to generate fixed or random characters and numbers, selectively replace data with null values, randomly substitute data from a list or column of values, or use SQL or regular expressions to transform data. You also have several advanced options to meet more complex business requirements, such as:

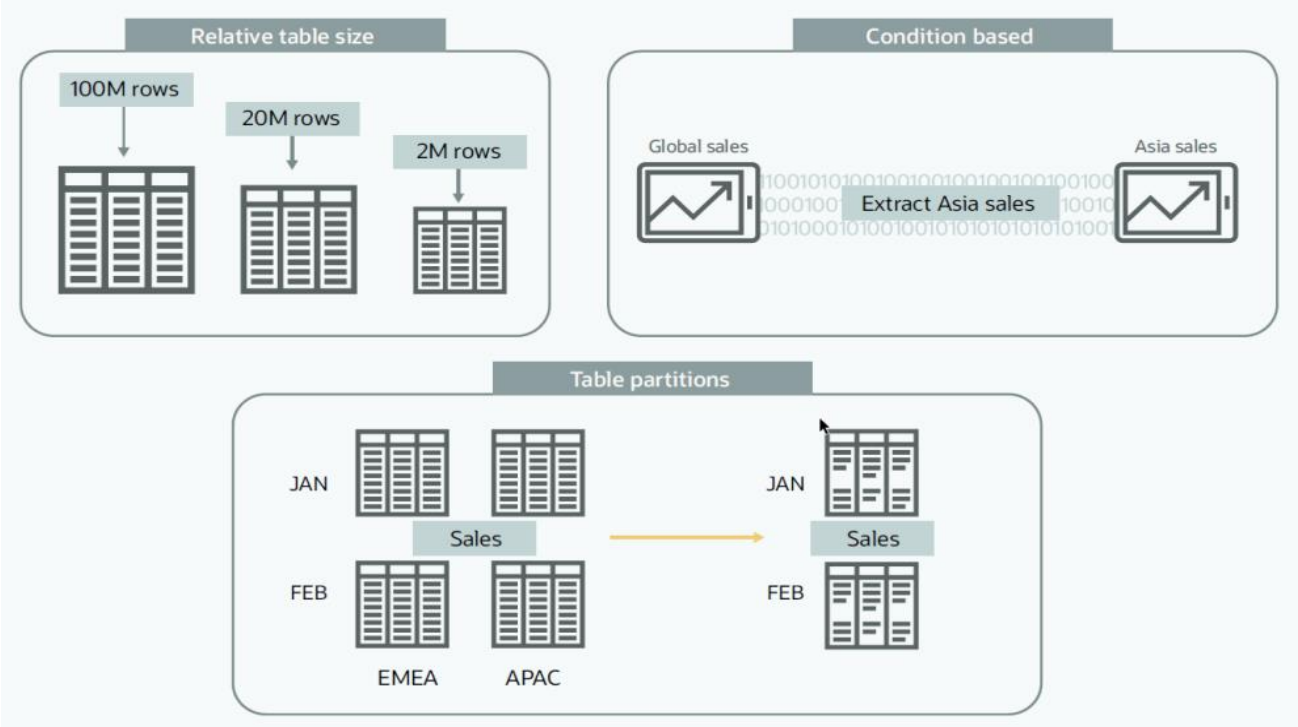
- **Shuffle masking** randomly shuffles data within a column. For example, columns containing salary information can be shuffled to break the employee-salary mapping.
- **Encryption** encrypts sensitive data using a cryptographic key while preserving the data's format. It's a reversible masking option, and you can decrypt your data using the same key. For example, reversible encryption is useful when masked data processed by a third party needs to be correlated with the source data.
- **Format Preserving Randomization** randomizes the data while preserving the input length, position, and case for characters and special characters.
- **Conditional masking** masks column data using different masking formats based on user-defined conditions. For example, in a column, the US identifiers can be masked using the Social Security Number format and the UK identifiers using the National Insurance Number format.

- **Compound masking** masks related columns as a group, ensuring the masked data across the related columns retain a consistent relationship. For example, compound masking of address fields such as city, state, and postal codes helps maintain consistency between these elements.
- **Deterministic masking** generates consistent masked output for a given input across application schemas and databases.
- **User-defined PL/SQL Masking** enables you to define custom masking logic or migrate your existing masking scripts.

DATA SUBSETTING

Data Subsetting helps reduce security risks and minimize storage costs by removing unnecessary data from a database before sharing it for non-production use. Data Masking and Subsetting provides goal-based and condition-based subsetting. A goal can be a relative table size, such as extracting a 1% subset of a table containing 10 billion rows. A condition can be based on factors such as time, such as discarding all user records created before a particular year. A condition can also be based on region, for example, extracting Asia Pacific information to support the development of a new application.

Figure 2. Data Subsetting Use Cases



CENTRALIZED ADMINISTRATION AND FLEXIBLE EXECUTION

Data Masking and Subsetting Pack installs by default with Oracle Enterprise Manager. It provides a centralized, unified, and browser-based GUI for administration. In addition to its intuitive GUI, Enterprise Manager provides Command Line Interface (EMCLI) to automate select data masking and subsetting tasks.

Masking and subsetting can be performed on a cloned copy of the original data, eliminating any overhead on production systems. Alternatively, it can be performed during database export, eliminating the need for staging servers. High-performance in-export masking is achieved through integration with the Oracle Database and Data Pump.

Data Masking and Subsetting can mask data in non-Oracle relational databases (MySQL, SQL Server, Sybase, DB2, Informix, and Teradata) by staging the data in an Oracle Database and leveraging the included Oracle Database Gateways.

SOFTWARE LIFECYCLE INTEGRATION

Data Masking and Subsetting is integrated with Oracle data management and testing tools. For example, Oracle Database Life Cycle Management Pack users can mask and clone databases in a single workflow. Oracle Data Integrator users can mask and subset data during data synchronization between source and target databases.

MASKING AND SUBSETTING FOR HYBRID CLOUD

Data Masking and Subsetting helps organizations achieve data privacy and compliance for non-production databases both on-premises and in the Oracle Cloud. For example, this hybrid management capability facilitates masking and subsetting while migrating data from on-premises to the Oracle Cloud.

MORE INFORMATION

For more information such as product FAQ, tutorials, documentation, customer references, and blog, please visit:

https://www.oracle.com/security/database-security/data-masking/

Related products

Oracle Database 23ai defense-in-depth solutions

- Oracle Advanced Security
- Oracle Key Vault
- Oracle Database Vault
- Oracle Label Security
- Oracle Audit Vault and Database Firewall

Connect with us

Call **+1.800.ORACLE1** or visit **oracle.com**. Outside North America, find your local office at: **oracle.com/contact**.

blogs.oracle.com facebook.com/oracle twitter.com/oracle