

# What Is a Data Inventory?

A **data inventory** lists datasets that are managed by an organization. Each listing is accompanied by information about the dataset, such as owner, frequency of update, and file format. (These details about a dataset are known as “metadata.”) For an organization, its data inventory is a foundational asset that can be used to prioritize work on various data projects, such as quality improvement and performance metric selection.

Organizations may have thousands of datasets across multiple servers, databases, and computers. The data inventory includes only those which are considered valuable to employees, departments, executive leadership, and the general public. These can range from individual datasets maintained by an individual, to tables from integrated databases.

## Why Conduct a Data Inventory?

There are many ways that a data inventory can improve an organization’s operations. Consider these benefits and which are most applicable to your unique needs and challenges.

- **Staff Collaboration:** When staff need information from another department, they use the inventory to learn what exists and how to find it. This enables collaboration on reporting and analytic projects.
- **Security:** Creating and managing a data inventory reduces risk and uncertainty by creating a checklist for security and compliance requirements. In some cases, the inventory can bring attention to data that the organization should no longer maintain.
- **Opening Data to the Public:** Having a complete inventory is also important when determining which datasets to release publicly. The data inventory can be used to prioritize the release of data.
- **Data Quality:** A data inventory prompts data owners to assess the quality of their data, and brings those quality ratings into a central location where the organization can prioritize efforts to improve data quality.
- **Performance Management:** Datasets in the inventory may contain individual data points that can serve as metrics for performance management. Having the datasets listed in an organized fashion facilitates conversations about metric selection.
- **Identifying Missing Data:** With a complete picture of existing data, an organization can identify what data does not exist, but would be useful for specific tasks.
- **Centralization and Coordination:** An inventory is a first step toward considering solutions for master data management and data warehousing.