

HOW (AND WHY) TO ENFORCE AND MEASURE DATA QUALITY



In today's information-driven world, implementing an effective data quality management or DQM strategy cannot be overlooked. DQM refers to a business principle that requires a combination of the right people, processes and technologies all with the common goal of improving the measures of data quality that matter most to an enterprise organization. Put more simply, effective DQM is critical for any enterprise business leader who wants to unlock accurate and actionable insights from their data sets.

But understanding the importance of DQM is one thing, enforcing data quality is another.

The Importance of Quality Data

With accurate data on your customers, you can better understand their needs and how best to reach them. There are a number of ways to capture data from your customers. Items like purchase information, feedback surveys and user profiles and credentials power sophisticated CRMs that just keep getting better at securely analyzing customer data.

But all of that means nothing if your organization can't separate the good data from the bad. Bad data can result in poor, or slow decision making from management, ineffective marketing and attrition of customers or users.

Cost Optimization

Overall, poor data quality is bad for business and has a significant cost as it relates to time and effort. In fact, [Gartner estimates](#) that the financial impact of the average financial impact of poor data quality on organizations is around \$15 million per year. Another study by Ovum indicates that poor data quality costs business at least [30% of revenues](#).

Effective Marketing

For marketer's, accurate, high-velocity data is critical to making choices about who to market to and how. This leads to better targeting and more effective marketing campaigns that reach the right demographics.

Better Decision-Making

A company is only as good as it's ability to make fast, accurate decisions. And in turn, this is driven by the inputs they are able to work with. The old adage "garbage in, garbage out" comes to mind in this regard.

The better the data quality, the more confident enterprise business leaders will be as it relates to mitigating risk in the outcomes and driving much more efficient decision-making.

Productivity

According to [Forrester](#), "Nearly one-third of analysts spend more than 40 percent of their time vetting and validating their analytics data before it can be used for strategic decision-making." Therefore, it stands to reason that when a data management process produces consistent, high-quality data more automation can occur. This means employees don't have to spend more time making tedious changes to data and can be more productive, in general.

Compliance

For many industries, housing data poses additional regulations and responsibilities with regards to compliance. Rigorous compliance concerns result in ongoing processes that must be performed routinely to get into compliance and remain that way.

Dashboard-type analytics stemming from good data have become an important way for organizations, like those in the financial industry, to understand, at a glance, if their organization is keeping up with compliance standards.

How to Enforce Data Quality

As mentioned above, [data quality management](#) is a principle in which all of a business' critical resources -- people, processes and technology -- work harmoniously to create good data. More specifically, data quality management, also known as DQM, is a set of processes designed to improve data quality with the goal of actionably achieving pre-defined business outcomes.

Data quality requires a foundation to be in place for optimal success. These core pillars include the following:

- The right organizational structure
- A defined standard for data quality
- Routine data profiling audits to ensure quality
- Data reporting and monitoring
- Processes for correcting errors in bad and incomplete data

Getting Started

If you are like many organizations, it's likely that you are just [getting settled in with big data](#). Here are our recommendations for implementing a strategy that focuses on data quality;

- **Assess Current Data Efforts** - an honest look at your current state of data management capabilities is necessary before moving forward.
- **Set Benchmarks for Data** - This will be the foundation of your new DQM practices. To set the right benchmarks, organizations must assess what's important to them. Is data being used to super-serve customers or to create a better user experience on the company website? First, determine business purposes for data and work backward from there.
- **Ensure Organizational Infrastructure is in Place** - Having the proper data management system means having the right minds in place who are up for the challenge of ensuring data quality. For many organizations, that means promoting employees or even adding new employees.

Roles and Responsibilities

An organization committed to ensuring their data is high quality should consider the following roles are a part of their data team:

- **DQM Program Manager:** This role sets the tone with regard to data quality and helps to establish data quality requirements. He or she is also responsible for keeping a handle on day-to-day data quality management tasks, ensuring the team is on schedule, within budget and meeting predetermined data quality standards.
- **Organization Change Manager:** This person is instrumental in the change management shift that occurs when data is used effectively, they make decisions about data infrastructure and processes.
- **Data Analyst I or Business Analyst:** This individual interprets and reports on data.
- **Data Steward:** The data steward is charged with managing data as a corporate asset.

Leverage Technology

Data quality solutions can make the process easier. Leveraging the right technology for an enterprise organization will increase efficiency and data quality for employees and end users.

How to Measure Data Quality

With the aforementioned information, you're ready to get started with data quality management. But if you already have processes for your data in place, you may be wondering how to measure it. That's what we will cover in the following:

Data Profiling

Data profiling is a good starting point for measuring your data. It's a straight-forward assessment that involves looking at each data object in your system and determining if it's complete and accurate.

This is often a preliminary measure for companies who use existing data but want to have a data quality management approach.

Data Quality Assessment Framework

A more intricate way to assess data is to do it with a Data Quality Assessment Framework known as DQAF. The DQAF process flow starts out like data profiling, but the data is measured against certain specific qualities of good data. These are:

- **Integrity:** how does the data stack up against pre-established data quality standards?
- **Completeness:** how much of the data has been acquired?
- **Validity:** does the data conform to the values of a given data set?
- **Uniqueness:** how often does a piece of data appear in a set?
- **Accuracy:** how accurate is the data?
- **Consistency:** in different datasets, does the same data hold the same value?

Using these core principles about good data as a baseline, data professionals can analyze data against their own real standards for each. For instance, a unit of data being evaluated for timeliness can be looked at in terms of the range of best to average delivery times within the organization.

Key Metrics

There are a few standardized ways to analyze data, as described above. But it's also important for organizations to come up with their own metrics with which to judge data quality. Here are some examples of data quality metrics:

1. Data-to-errors ratio: analyzes the number of errors in a data set taking into account its size.
2. Empty values: this is an assessment of how much of the data set contains empty values.
3. Percentage of "dark data": dark data is a term that means unusable data. The more of that in a set the worst off it is.
4. Email bounce rates: if your emails aren't going through at a higher percentage than what is typical. You could have a data issue.
5. Time-to-value: this is a ratio that represents how long it takes you to use and access important data after input into the system. It can tell you if data being entered is useful.

Final Thoughts

In a technologically-driven world that seeks increased personalization from enterprise organizations, data is the key to success. To be effective with data you have to shift your mindset from simply capturing it, to implementing strategies to ensure its quality.

This means having the right people, processes and technology in place to do so.