Poor Data Health is Costly but Improving is Easier Than Ever

By E. Wallace

Potential data insights and what companies can actually get from their data are two very different things. In fact, one frightening study suggested that more than 70% of data goes unused for decision making. Several obstacles to data insight include the wrong tools, poor infrastructure, or lack of expertise. Another significant challenge to data-driven decision making is data health. Companies must take charge of data health for full digital transformation, and luckily, it's now easier than ever. Let's explore why.

What is Data Health?

Data health is a company's data-readiness. To determine data readiness, businesses should consider if data is valid and of sufficient quality. But what does quality mean in the context of business data?

Each company may have a different set of quality metrics for determining if data quality is sufficient, but a few common benchmarks are:

- **Is data complete?** The datasets must contain all the necessary information to answer the query.
- **Is the accuracy trustworthy?** Certified data, a clear lineage, reliable sourcing — all these things provide trust for accurate data.
- **Is each entry unique?** Accidental repeat entries throw off data accuracy and wreak havoc with final insights. In some cases this is a consistency issue, but in others, multiple entries can be the result of acquiring.
- **Is this data timely?** There may be reason to check historical data, but in many cases, data should be the most up to date option for the query.

© 2021 The Modern Data Company. All trademarks are properties of their respective owners.

Without data quality, businesses cannot make decisions and risk major compliance issues. Therefore, regular data health checks should be on the schedule.

Human error causes a significant number of data incidents, both accidental and security breaches. Investing in regular data health checks is not just an ideal; it's an essential part of running a data-driven organization.

- **Does the dataset follow a consistent pattern?** For enterprises with large and varied datasets, inconsistencies within and between sets can wreck accuracy.

- **Can the right people access it?** Part of quality means the data is retrievable for those who need it. That means it's not locked away or stored in an impossible framework that fragments it. At the same time, companies must be sure that data is secure from hackers or even competitors.

The Cost of Ignoring Data Health

Poor data quality costs each business millions per year on average, according to a survey from Gartner published back in 2017. For industries with sensitive data like healthcare, those bad data can be particularly egregious. A 2021 study published in the HIPAA Journal noted an average of two data breaches per day in healthcare.

Even without data leaks putting companies at risk for serious sanctions, bad data costs companies in other ways. Decision-makers, IT teams, and departmental groups wrestle with poor data quality in their day-to-day decisions, creating bottlenecks for ingestion and scrubbing and wasting time on poorly executed business strategies backed by faulty data. The end result is over $3 trillion per year in soft and hard costs, thanks to poor data health. Without a plan, data moves from asset to liability quickly.

Developing a Plan for Data Health

Businesses must recognize their data as a whole, almost living being. As such, maintaining data health is every bit as important as maintaining a person's health.

Decision-ready data only happens with a plan. That plan includes three essential components.

- **Proactive** - Identifying faulty data elements and security concerns before it ever becomes an issue. This starts at the most granular level.

- **Reactive** - Red flags in the data and actual incidents require swift, resolute action.

- **Support** - A data culture with a clear chain of responsibility provides a foundation for monitoring as well as data management.

These components provide the foundation for managing data and ensuring that the quality of all data, both real-time and historical, is high enough to get to valuable insights. A troubling statistic back in 2017 claimed that only around 3% of companies' data met basic quality standards. In the years since then, things have only moderately improved.

© 2021 The Modern Data Company. All trademarks are properties of their respective owners.

A data fabric provides a unified environment, allowing all services, technologies, and data storage to run on the same architecture. It streamlines management and offers a revolutionary way to manage governance.

Leveraging a Data Fabric to Improve Data Health

Data fabric offers:

- **Connection:** Companies can easily link to any data source, ensuring that all data follows consistent protocols.
- **Ingestion:** As data comes in, a data fabric ensures that data remains complete, consistent, and available for those who need it.
- **Management for multiple environments:** On-premise, cloud, and hybrid, whatever companies need to ensure that data remains timely and secure; a data fabric offers integration for all types.
- **Streamlined sharing:** With granular governance capabilities and API support, everyone who needs the data gets it.

Data fabrics manage everything from collection to sharing to insights. It addresses challenges like unifying multiple data environments and reducing the reliance companies have on legacy systems.

By bringing everything under one roof, businesses can manage, scale, and future-proof data.

DataOS® is a first-of-its-kind data operating system. It's designed to remove complexity so that businesses can manage data health easily with fewer mistakes. It offers declarative data management with AI-powered determination for the most appropriate dataset and transforms data security through a one-of-a-kind tag-based system. It's time to see what real data health can be.

© 2021 The Modern Data Company. All trademarks are properties of their respective owners.