

Data quality is crucial – it assesses whether information can serve its purpose in a particular context (such as data analysis, for example). So, how do you determine the quality of a given set of information? There are data quality characteristics of which you should be aware.

There are five traits that you'll find within data quality: accuracy, completeness, reliability, relevance, and timeliness

Accuracy

As the name implies, this data quality characteristic means that information is correct. To determine whether data is accurate or not, ask yourself if the information reflects a real-world situation. For example, in the realm of financial services, does a customer really have \$1 million in his bank account?

Accuracy is a crucial data quality characteristic because inaccurate information can cause significant problems with severe consequences. We'll use the example above – if there's an error in a customer's bank account, it could be because someone accessed it without his knowledge.

Completeness

"Completeness" refers to how comprehensive the information is. When looking at data completeness, think about whether all of the data you need is available; you might need a customer's first and last name, but the middle initial may be optional.

Why does completeness matter as a data quality characteristic? If information is incomplete, it might be unusable. Let's say you're sending a mailing out. You need a customer's last name to ensure the mail goes to the right address – without it, the data is incomplete.

Reliability

In the realm of data quality characteristics, reliability means that a piece of information doesn't contradict another piece of information in a different source or system. We'll use an example from the healthcare field; if a patient's birthday is January 1, 1970 in one system, yet it's June 13, 1973 in another, the information is unreliable.

Reliability is a vital data quality characteristic. When pieces of information contradict themselves, you can't trust the data. You could make a mistake that could cost your firm money and reputational damage.

Relevance

When you're looking at data quality characteristics, relevance comes into play because there has to be a good reason as to why you're collecting this information in the first place. You must consider whether you really need this information, or whether you're collecting it just for the sake of it.

Why does relevance matter as a data quality characteristic? If you're gathering irrelevant information, you're wasting time as well as money. Your analyses won't be as valuable.

Timeliness

Timeliness, as the name implies, refers to how up to date information is. If it was gathered in the past hour, then it's timely – unless new information has come in that renders previous information useless.

The timeliness of information is an important data quality characteristic, because information that isn't timely can lead to people making the wrong decisions. In turn, that costs organizations time, money, and reputational damage.

“Timeliness is an important data quality characteristic – out-of-date information costs companies time and money”

In today's business environment, data quality characteristics ensure that you get the most out of your information. When your information doesn't meet these standards, it isn't valuable. Precisely provides data quality solutions to improve the accuracy, completeness, reliability, relevance, and timeliness of your data.