What is data quality?

Public sector organisations need the right data in order to run good services, make the right decisions, and create effective policies. Data takes many forms including budget figures, employee data, survey responses, and data gathered in digital services. Everyone has a role to play in ensuring data quality. You can read more about the importance of good quality data in the [Government Data Quality Framework](#).

In this article we're going to look at what data quality is, and the different paths towards good quality.

What does good quality look like?

Good quality data is data that is fit for purpose. That means the data needs to be good enough to support the outcomes it is being used for. Data values should be right, but there are other factors that help ensure data meets the needs of its users.

Quality of data content

Having good quality data does not mean every value must be perfect; good quality will be different for different data sets. Quality can be measured using [six dimensions](#): completeness, uniqueness, consistency, timeliness, validity and accuracy. Different data uses will need different combinations of these dimensions; there are no universal criteria for good quality data. It is important to actively manage quality, and work to improve poor quality.

Quality of data processes

Having good quality data is a great start but it needs to be maintained. When data is moved or changed there is a chance for quality problems to be introduced. Automating manual data processes together with robust validation rules can prevent errors and improve consistency.

Documenting the processes used in your data's journey improves the organisation's understanding and helps to ensure consistency when handling the data.

Quality of data sets

If we start with values that are right, and process them well, we also need to ensure that the data we collate, package and share as a data set is good quality. Providing data in an agreed format or specification ensures consistency and makes it easier for users to process and analyse further.

All datasets should have metadata — information about the data that helps

people understand what it is (and is not!).

Quality of analysis

Having the right data values, good processes and well created datasets gives us the best foundations for analysis. Continuing quality assurance throughout the analytical process helps ensure quality analysis; see the [Aqua Book](#) for more detail. When analysis needs to be delivered under significant time constraints, it may not be possible to carry out full quality assurance checks. Read the [guidance on urgent quality assurance](#) for more information.

Supporting good data quality

These aspects of quality don't stand alone. Other things that can support an organisation in achieving good quality data include:

Data governance

Clear roles, responsibilities, policies, principles, and organisational structures ensure data is managed well, in a way that benefits the whole organisation.

Design

Good design of services, data architecture and data collection will build the best foundations for good data quality.

Data Management

Having the right activities and processes around data ensures it is properly managed in an organisation. These can cover quality, design, sharing, access, security and many other aspects of data. Look out for forthcoming guidance on a single data maturity model for use across government. This will help you assess your current data management practices.

Reproducible Processing

Automating data processes reduces errors. The Office for Statistics Regulation recommends [Reproducible Analytical Pipelines](#) as the best approach to use for official statistics.

Metadata

Metadata is information about data such as a description of the data source, its purpose and processing. Relevant metadata supports data quality work and helps users to assess whether the data set is adequate for their use. For more guidance, see the [Metadata Standards](#) published by the Data Standards Authority.

Standards

Standards define consistent ways of capturing or storing data. Standards

improve sharing and integration of data. The [Data Standards Authority](#) leads the cross-government conversation around data standards. Standards help to establish consistency and validity of data; they do not guarantee data quality. If you are designing services to collect data from users, refer to the [gov.uk design patterns](#).

Records Management

Good organisational policies and procedures for storing and retaining information can grow into good practice for your data. Good records management depends on clear ownership, maintenance of assets, retention and disposal, audit trails and metadata, just like good data management.

How can data quality be achieved?

Good quality means:

- good design
- having the right values in your data
- processing that data well
- forming it into good quality datasets accompanied by metadata
- analysing the data properly

This will be easier to achieve if you have the right governance, information and data management, and standards in place.

The Government Data Quality Hub (DQHub) is developing tools, guidance and training to help you with your data quality initiatives. You can find the [Government Data Quality Framework](#), tools and case studies on [our website](#).

We can also offer tailored advice and support. Please contact us by emailing dqhub@ons.gov.uk.