Data Quality Management Policy

1. Data Quality Policy

Implemented on	August 2023
Last review on	15th August 2023
Next review due on	15th August 2024
Approved by	Data Governance Committee
Policy owner (name)	Laura Payne
Policy owner (division)	Data Quality Hub
Main point of contact	dqhub@ons.gov.uk
Published version link	

Index

Scope

This policy covers all the data that the Office for National Statistics (ONS) holds either as a data controller or a data processor, including those obtained through surveys and from administrative sources in the private or public sectors. The policy is valid from data acquisition and collection by the ONS, and it also extends through the lifecycle up to and including data archiving and data disposal. This does not just include data sources we acquire/collect and the outputs we produce, but also any intermediate data

product necessary to the data lifecycle. This data includes data that is not collected or used for statistics and statistical research such as HR and finance data.

This policy applies to all UK Statistics Authority including ONS employees, including staff on fixed-term, temporary or permanent contract, staff on secondment, students, and contingent workers. Any user of ONS data (internal or external) must be encouraged to raise any issues related to data quality.

Background

This policy has been created to ensure that the quality of data is considered from the start, understood, assessed, communicated, and managed consistently throughout its full data lifecycle as highlighted by the data quality pillar within the ONS Data Strategy.

The Our Data Strategy specifies our commitments for managing data, and the data policies support the implementation of these commitments. This data quality policy supports the delivery of the ONS Data Strategy, and specifically the implementation of the Data Quality Pillar, as well as setting the requirement for the management and governance of data quality in the ONS. This policy is in support of the Code of Practice for Statistics, specifically its Quality pillar, which provides guidance to ensure that data and methods that produce assured statistics are adhered to.

It is important to ONS to establish and maintain a consistent approach to assuring data quality that covers the Data Management Association (DAMA) Data Quality Dimensions. This is to support the overall strategic direction and achieve corporate objectives by using policies and data standards to uncover and address the issues and risks associated with the quality of the ONS data.

Finally, this policy covers the full data lifecycle and should be used in conjunction with the Government Data Quality Framework.

Policy Statement

Quality is a cornerstone of work at the ONS. We have a long history of ensuring quality in our statistical work, but trustworthy statistics must be built on data that are fit for purpose. We can and must do more to ensure that we better understand the quality of all data we hold and use to demonstrate that we are a trustworthy organization. We must ensure that we better understand the quality of all data we use, collect, hold, and produce. ONS must establish a culture that puts data quality at the core of every activity and remember that quality begins at the beginning of the data lifecycle. By ensuring that data quality is assessed at the earliest stages of planning for services,

data storage, design, and data collection we can enhance the trustworthiness of the data that we hold and use and build public confidence in our statistics.

Good quality data is data that is fit for purpose: the data needs to be good enough to support the outcomes it is being used for, and as such "the level of quality required will vary depending on the purpose. Data quality is more than just data cleaning." [1]This policy demonstrates the need for data quality to be governed, communicated and that data quality should be at the core of every activity.

Policy Detail

1. Data quality shall be governed across the data lifecycle to ensure that data quality is appropriately managed:
- Processes shall be in place for systematically reviewing and assessing the quality of data
- Informed decision making shall be considered, if the quality is low then you must make sure that you/others are aware of the consequences of the risk
- Data Quality shall be considered/built in from the start. It must not be an consideration only at the end of the data lifecycle , i.e. , do not depend on data cleansing or editing the data
- Any remediation activities shall be documented to include the impacts to the rest of the data lifecycle
- Data quality issues shall be identified, documented, escalated and communicated; this is to include any third parties we share data with
- Processes shall be in place for reviewing the quality of data, this will be contained in the Data Quality Standard (TBC – Owner)
- Substantial changes to the Data Quality Management Policy shall be approved by the Data Governance Committee

2. Data quality shall be communicated to all users and data quality issues shall be documented at all stages of the data lifecycle, risks shall be known and understood:
- Any changes or remediation of data quality shall be communicated to all users within the data life cycle
- Data product users shall be able to see what inputs went into it and have access to information on the quality of those sources
- All supporting data quality information shall be communicated, kept up to date, and its impact must be clearly explained to users. It shall inform the users through tailored communication to make their own opinion of the quality of the data based on their specific needs.

---

[1] The Government Data Quality Framework
https://www.gov.uk/government/publications/the-government-data-quality-framework

- All contracts/agreements for supply or use of data shall consider how to maintain data quality, assess it, communicate it, and escalate any issues
- ONS must have in place robust and transparent procedures to capture feedback on data quality and resolve quality challenges by taking appropriate and proportionate actions
- A process that details the path to engagement and resolution shall be in place, process owner TBC

3. All teams in ONS shall establish a culture of good data quality: data quality shall be at the core of every activity:
- Data quality will be part of all data management policies signifying its importance across ONS
- Every member of staff should be aware of policies relating to data quality and related policies (e.g., data protection).
- All ONS staff shall be empowered to challenge data providers where data isn't fit for purpose. There shall be access to data quality information that is required so the users can make informed decisions
- Data Quality culture shall work in both directions so that information can be shared throughout the data life cycle, questions and challenges shall be encouraged

Integrated Data Service

The above policy detail is also applicable to The Integrated Data Service.

The Integrated Data Service will provide data across government for many uses, the quality of this data will be important to ensure that outputs are of a minimum standard:
- Data quality management shall be clearly defined and owned
- Guidance shall be made available to users so that they are aware of potential escalation should data quality problems occur
- Users of IDS shall report data quality concerns so that they can be dealt with appropriately
- Quality information/metadata shall be available at the same time of the data, regularly updated when necessary and clearly communicated to users.
- Inappropriate use of data shall be challenged. The Code of Practice for Statistics states that statistics should be based on data sources that are appropriate for the intended uses.

- The producer of the data shall do everything that they can to minimize the risk of misuse or misinterpretation.

Roles and responsibilities

ONS staff working with data
- Compliance to the Data Quality Policy.
- Aware of their individual responsibilities contained within this policy relating to data quality and quality assurance of data and their impact upon the quality of data, statistics, and analysis
- Aware of their obligations under the ONS Statistical Quality Improvement Strategy to "ensure our data are of sufficient quality and communicate the quality implications to users"
- Awareness of the Quality Assurance of Administrative Data (QAAD) standard supporting the Code of Practice for Statistics principle on Q1: Suitable data sources
- Raise concerns regarding data quality through line management and Quality Champions, and through the Statistical Quality Maturity Model Exercise.
- Ensure that risks and issues associated with data quality, together with associated actions, are appropriately recorded in their Divisional Quality Improvement Plan
- Ensure that the commentary accompanying statistics and supporting Quality and Methodology Information appropriately reflect the quality of the data

Line Manager responsibilities
- Awareness of this Data Quality Policy as well as other relevant policies/strategies in their specific areas.
- Monitor, challenge and assure that policy is being followed within their management.
- Support their staff in implementing processes that promote good data quality practices
- Consider the quality of data before them when making or advising on decisions and ensuring that they have the information required.

- Collaborate across different business areas to ensure that data quality issues impacting are addressed throughout the data lifecycle and not strictly within business areas.
- Records of downstream users of any data products should be maintained so that it is clear who data quality information needs to be shared with

Data Quality Hub
Accountable to the Chief Data Officer and responsible for:
- Providing expert advice and guidance on what is best practice related to data quality
- Act as a ONS-wide sponsor for data quality and clearly communicate and coordinate data quality requirements
- Supporting data quality remediation and escalation to the Quality Committee as and when required.

Quality Committee
Accountable to the National Statistician and the UKSA Board and responsible for:
- Approval of complex data quality remediation as and when required
- Mitigating tactical risks and assessing the mitigating actions for incidents related to data quality.

UK Statistics Authority – Data Governance Legislation and Policy team
Accountable to the Data Governance Committee and responsible for:
- Providing independent scrutiny and assurance against the policy on behalf of the Data Governance Committee
- Support data policy development to ensure that data quality is well represented.
- Provide independent scrutiny of data quality audits, as and when required.
- Ensuring that data quality activities remain transparent and align with this policy.

Data Governance Committee
Accountable to the National Statistician and UKSA Board and responsible for:
- Monitoring and reviewing how this policy is implemented, acting on the assurance and maturity assessment information provided to them.
- Mitigating strategic risks and assessing the mitigating actions for incidents related to data quality.

Supporting Documents

Data Quality Framework

Strategic Links

- CDDO 2025 Commitment
- Our Data Strategy
- Joined up data in government: the future of data linking methods
- Digital Economy Act Accredited Processing Environments
- UKSA Governance Guide
- UKSA Quality Committee Terms of Reference
- Data Governance Committee Terms of Reference (TBC)
- https://www.ons.gov.uk/aboutus/transparencyandgovernance/datastrategy/relevantlegislation

Data Policies

- Metadata Policy
- Data Acquisition Policy
- Data Ethics Policy - Office for National Statistics (ons.gov.uk)

Data Standards

- ONS Data Standards
- ONS Data Principles
- Ethical Principles – UK Statistics Authority

Survey Supporting documents

- Survey strategy
- Survey playbook
- https://analysisfunction.civilservice.gov.uk/wp-content/uploads/2018/02/Generic-Statistical-Business-Process-Model.pdf

Admin data:

- UK: Office for Statistics Regulation – Administrative Data and Official Statistics
- Quality Assurance of Administrative Data (QAAD)
- ONS: QAAD

Legislation

- UK General Data Protection Regulation principles
- Data Protection Act 2018 (link TBC)
- Statistics and Registration Service Act 2007(link TBC)

Other useful links:

- Jargon Buster
- ONS Statistical Quality Improvement Strategy
- Code of Practice for Statistics – Quality
- ONS Quality Questions and Red Flags
- ONS Business Glossary (TBC)

2. **Appendix**

2.1 Core Data Quality Dimensions

Core Data Quality Dimension	Explanation
Completeness	Completeness describes the degree to which records are present. For a data set to be complete, all records are included, and the most important data is present in those records. This means that the data set contains all the records that it should and all essential values in a record are populated. It is important not to confuse the completeness of data with its accuracy. A complete data set may have incorrect values in fields, making it less accurate.
Uniqueness	Uniqueness describes the degree to which there is no duplication in records. This means that the data contains only one record for each entity it represents, and each value is stored once.

	Some fields, such as National Insurance number, should be unique. Some data is less likely to be unique, for example geographical data such as town of birth.
Consistency	Consistency describes the degree to which values in a data set do not contradict other values representing the same entity. For example, a mother's date of birth should be before her child's. Data is consistent if it doesn't contradict data in another data set. For example, if the date of birth recorded for the same person in two different data sets is the same.
Timeliness	Timeliness describes the degree to which the data is an accurate reflection of the period that they represent, and that the data and its values are up to date. Some data, such as date of birth, may stay the same whereas some, such as income, may not. Data is timely if the time lag between collection and availability is appropriate for the intended use.
Validity	Validity describes the degree to which the data is in the range and format expected. For example, date of birth does not exceed the present day and is within a reasonable range. Valid data is stored in a data set in the appropriate format for that type of data. For example, a date of birth is stored in a date format rather than in plain text.
Accuracy	Accuracy describes the degree to which data matches reality.

	Bias in data may impact accuracy. When data is biased it means that it is not representative of the entire population. Account for bias in your measurements if possible, and make sure that data bias is communicated to your users. In a data set, individual records can be measured for accuracy, or the whole data set can be measured. Which you choose to do should depend on the purpose of the data and your business needs.