

Quality Questions and Quality Red Flags

This document is published and maintained by the Data Quality Hub and the Analysis Standards and Pipelines Hub. It aims to provide support to colleagues working on statistical outputs and analysis in ONS.

The tab 'Quality Questions' include a set of questions that the analyst can use to interrogate their work with their team, and will support them in assuring the quality of their work. The document explains why the question is important, and what help, guidance, and support is available.

The tab 'Quality Red Flags' include a set of statements that can help the analyst and their team to identify potential risks to quality that might benefit from further look into. These statements are not meant to be considered as a value judgment, but more as a suggestion that best practice and support could be beneficial.

If you have any questions or feedback about this document, please reach out to the [Data Quality Hub](#) or the [Analysis Standards and Pipelines Hub](#). We will keep this updated and we always welcome any input from anyone working on statistical releases or analysis to ensure this document can be the most helpful possible.

Last updated 30/09/2022

Quality Questions

Quality Red Flags

Quality Question	Why do I need to know the answer to this?	What help is available here?
What is the need for this analysis or statistical release?	Understanding why the analysis/statistical release is needed and what it will be used for is critical for understanding whether what you have done is fit for purpose. If you are responsible for part of an analytical/statistical process, understanding the end use will help you to make sure that your part does what is needed to meet user needs.	Guidance: The AQUA Book Guidance: Analysis Functional Standard Get in touch with the User Research team
Who uses your analysis or statistical release?	Understanding who uses your analysis/statistical release will help you to make sure that it meets their needs. It also helps you to tailor your outputs to make sure all your users are fully supported in using the outputs effectively.	Guidance: User engagement top tips Get in touch with the User Research team and the External Affairs team
What analytical question you are addressing?	Having a clear understanding of the problem your team is trying to solve ensures that the analysis you design is fit for purpose. If you do not know how your work is contributing to answering an analytical need, you may be unaware of important requirements or limitations for your part of the work.	Guidance: The AQUA Book Guidance: Analysis Functional Standard Get in touch with the Analysis Standards and Pipelines Hub The Government Data Quality Framework Introduction to data quality Introduction to data quality assessments Tips for urgent quality assurance of data Quality Assurance of Administrative Data (QAAD) toolkit Quality of Admin Data in Statistics (Draft guidance) Data Quality Action Plans Data Quality Dimensions Quality Assurance: Four Areas of Practice Get in touch with the Data Quality Hub
What is the quality of the data that you use?	Understanding the quality of your data inputs is critical. It enables you to assess limitations and uncertainty in the inputs and how they feed through to your outputs. If you don't understand this, you will be unable to assess the quality of your process or your outputs. Understanding the quality of the data will allow you to assess whether the data can be used to address your analytical questions and the underlying user needs.	

Quality Question	Why do I need to know the answer to this?	What help is available here?
How did you choose the methods for the analysis or statistical release?	You should be able to explain why you chose the method (or set of methods) that you are using to produce your analysis or statistical release. A clear rationale for your method gives you and your users confidence that your choice is based on sound reasoning and evidence.	<p>Guidance: The AQUA Book Analysis Functional Standard</p> <p>Get in touch with the Analysis Standards and Pipelines Hub and the Methodology Advisory Service (MAS)</p>
How do you know the method you are using is appropriate?	You should be able to explain why the method(s) you use are suitable for this analysis/statistical release and be able to support your choice with evidence. This might include reference to academic peer review or other projects that are similar. If you can't explain why you chose the methods you use and why they are right for your analysis and the data you are using, you cannot be sure that your approach is sound.	<p>Guidance: The AQUA Book</p> <p>Get in touch with the Methodology Advisory Service (MAS)</p>
Can you summarise and explain the end-to-end process of your analysis or statistical release for somebody who asks about it?	Having an overview of your analysis/statistical release (especially if you only work on part of it) ensures that you and your team understand how your work feeds into the wider product. It can help you to identify potential quality risks or issues, both upstream and downstream of your own work as well as how your activity supports and underpins downstream processing.	<p>Guidance: The AQUA Book</p> <p>Generic Statistical Business Process Model (GSBPM)</p>
How do you know that your analysis or statistical process is working correctly?	You need to be sure that your analysis produces the outputs that you think it should and that the processes you run work as expected. If you cannot demonstrate that scripts and processes you have set up are functioning correctly, you cannot confirm the quality of the results.	<p>Guidance: Verification and validation for the AQUA Book</p> <p>Get in touch with the Analysis Standards and Pipelines Hub</p>



Quality Question	Why do I need to know the answer to this?	What help is available here?
Would another analyst be able to pick up from where you left off and reproduce or continue the work (without talking to you first)?	Your analysis must be well documented so that somebody new can understand it and pick it up. Poor documentation means that other people will not understand why the process is configured as it is, how the process works or how to run the process safely - potentially leading to errors.	<p>Guidance: The AQUA Book</p> <p>Guidance: QA of Code for Analysis and Research</p> <p>Get in touch with the Analysis Standards and Pipelines Hub</p> <p>Guidance: Analysis Functional Standard</p>
If you find a mistake in your analysis, do you have a clear and efficient process for addressing the issue and preventing it from happening again?	Analysis with lots of manual steps or that uses several tools is usually hard to assure. When problems happen, finding out how and why can be really difficult, and this does not apply only to manual processing. If your process is in this category, you are probably carrying extra quality risks.	<p>Guidance: QA of Code for Analysis and Research</p> <p>Guidance: The AQUA Book</p> <p>Get in touch with the Statistics Head of Profession Office and/or the Quality Champions Network</p>
Do you consistently use peer review to check scripts and code, documentation, implementation of methods, processes and outputs?	Peer review is a standard part of analysis best practice. It is helpful because it helps to identify where steps are unclear, documents are hard to understand or there might be problems with calculations or implementation of methods. Routine peer review helps to improve the quality of processes and to reduce risk by identifying potential problems.	<p>Guidance: Quality assurance of code for analysis and research</p> <p>Guidance: Quality statistics in government</p> <p>Get in touch with the Analysis Standards and Pipelines Hub (scripts, code, and documentation, peer review of models), the Data Quality Hub (processes and outputs), and the Methodology Advisory Service (methods).</p>
What are the limitations of your analysis or statistical release?	You should be able to explain any issues or limitations with your analysis/statistical release, and how they impact on potential use. A formal log of issues and limitations is a good way to make sure everybody in the team understands potential problems.	<p>Guidance: The AQUA Book</p> <p>Get in touch with the Analysis Standards and Pipelines Hub</p>



Quality Question	Why do I need to know the answer to this?	What help is available here?
Could you give a clear account of what can and cannot be inferred from your analysis or statistical release?	Ultimately our analysis/statistical releases should inform public commentary and decisions taken. A clear statement of the extent to which the analysis does and does not support these will help reduce the chance of errors in these actions.	Guidance: Communicating quality, uncertainty and change e-learning: Communicating quality, uncertainty, and change Guidance: The AQUA Book Get in touch with the Analysis Standards Hub and Pipelines Hub and the Methodology Advisory Service (MAS) . Guidance: Communicating quality, uncertainty and change e-learning: Communicating quality, uncertainty, and change Guidance: The AQUA Book Get in touch with the Data Quality Hub and the Methodology Advisory Service (MAS)
Have you assessed the impact of the limitations and set out how they will affect the quality and use of the outputs?	Where you have identified limitations (for example data quality issues) you should be able to explain how they impact on the quality of the analysis. If you cannot, you do not know how good the output is.	Guidance: Communicating quality, uncertainty and change e-learning: Communicating quality, uncertainty, and change Guidance: The AQUA Book Get in touch with the Data Quality Hub and the Methodology Advisory Service (MAS)
How do you measure and report uncertainty in your analysis or statistical release?	No analysis is perfect and no data are completely correct. You should be able to explain how you have assessed and measured the uncertainty that affects your analysis.	Guidance: Communicating quality, uncertainty and change e-learning: Communicating quality, uncertainty, and change Guidance: The AQUA Book Get in touch with the Analysis Standards and Pipelines Hub and the Data Quality Hub Guidance: Quality statistics in government Tips for urgent quality assurance of ad-hoc analysis Guidelines for measuring statistical quality Mandatory training on Quality Statistics in government Mandatory introductory training on Code of Practice
What is the assessment of the quality of your analytical outputs?	Understanding and reporting on the quality of your analytical outputs is critical to ensure fitness for purpose, as well as trust and reliability. This ensures that analysis can appropriately inform decision-making. Your quality assessments are also key information to be shared with your users (for instance, through the Quality and Methodology Information Reports - QMIs) as well as being a requirement of the Code of Practice for Statistics.	For support and advice, get in touch with the Data Quality Hub

[For more Best Practice Resources and Guidance, check our list.](#)



Quality Red Flags	Why does this matter?	What help is available here?
I don't know who the Quality Champion(s) for my division is	Your local Quality Champion can advise on best practice for assuring your analysis, and will have a wider perspective on the work of your division. Make use of their knowledge and experience. New members are always welcome!	The Data Quality Hub can advise on who the Quality Champions are in your area.
I am not sure what the best-practice guidance is for my work.	If you don't know the recommended way to do things, you are unlikely to follow best practice by chance. Having a good understanding of best practice will help you to improve quality and reduce risk.	The Analysis Function guidance hub has helpful guidance about analytical best practice. The Data Quality Hub , the Analysis Standards and Pipelines Hub , and the Methodology Advisory Service can support and share best practice guidance.
I don't know who to contact about the methods I use.	Things change - society and the economy, households and businesses evolve and so do our data. We need to adapt methods to address these changes and sustain the quality of our outputs. Lack of contact with methodologists and academics may indicate a deficit in the quality of methods used.	The Methodology Advisory Service and the Analysis Standards and Pipelines Hub can help you with methods and their implementation. The Government Data Quality Framework Introduction to data quality Introduction to data quality assessments Tips for urgent quality assurance of data Quality Assurance of Administrative Data (QAAD) toolkit Quality of Admin Data in Statistics (Draft guidance) Data Quality Action Plans Data Quality Dimensions Quality Assurance: Four Areas of Practice
I don't know how to understand the quality of my data and its implications on my results	If you don't know how to assess the quality of the data used in your analysis you will be unable to measure the quality of the outputs or to decide if they are fit for purpose.	The Data Quality Hub can advise and provide support on assessing data quality and understanding implications for results and analysis

Quality Red Flags	Why does this matter?	What help is available here?
I don't know how my outputs will be used	If you don't know what your outputs will be used for you cannot be sure that they will meet user needs. A good understanding of likely use cases is an essential part of making sure your analysis is fit for purpose.	Guidance: User engagement top tips Your team leader, Deputy Director or User Engagement champion can support in understanding how your outputs will be used.
I can't describe the assumptions of my analysis or statistical release, when they were made and who made them and signed them off	All analysis involves assumptions. Understanding what those are (and being clear how and why you made them) is critical for understanding why the analysis works the way it does. Keeping a log of the assumptions you make, when you made them and who agreed them is a really helpful way to keep track of this.	The Analysis Standards and Pipelines Hub can advise you on how to track and manage your assumptions.
My colleagues never challenge me about the results I produce.	People working with your results will have expectations about what the results will be and observed departures from these expectations are an important part of the QA process. An independent layer of QA, possibly including a curiosity-type session will help with this assurance. This can be on the overall results or a component, e.g for a sector or for the latest set of weights.	The Data Quality Hub and the Analysis Standards and Pipelines Hub can support you in developing further QA and/or start curiosity sessions.
I can't describe how important decisions were made about my analysis or statistical release, when they were made or who made them and signed them off	All analysis involves decisions. Understanding what those are (and being clear how and why you made them) is critical for understanding why the analysis works the way it does. Keeping a log of the decisions you make, when you made them and who agreed them is a really helpful way to keep track of this.	The Analysis Standards and Pipelines Hub can advise you on how to track and manage your decisions.
I don't fully understand the end-to-end process of the analysis or statistical release	It is important that you understand how your work feeds into the bigger picture, especially if you only work on a small part of an analysis workflow. If you are not aware of issues with the inputs to your work, or how your work feeds into other work later in the process you may miss important quality issues or fail to include important quality checks that don't impact you directly but have a big effect later on in the process.	Your team leader or Deputy Director can support in understanding the end-to-end process of the analysis.



Quality Red Flags	Why does this matter?	What help is available here?
I can't explain how the work I do impacts on downstream processes	Most analysis feeds into other work. Your outputs might be re-used by analysts in a policy department, for example, or to add value to another ONS output. While it is not possible to keep track of every possible use of your data, you should have a good idea of the main uses that your analysis supports so you can be sure it meets user needs.	Guidance: User engagement top tips Generic Statistical Business Process Model (GSBPM) Your team leader or Deputy Director can support in understanding the impact of your work and other processes dependent on it.
It is difficult to assess the errors and uncertainty in the analysis or statistical release	You should be able to measure and explain how uncertainty affects your analysis. What is the margin of error around your outputs, for example? If it is difficult to measure uncertainty you should think about what this means for use of the outputs and whether there is anything you can do to improve your assessment of uncertainty. No analysis is 100 percent certain!	The Data Quality Hub and the Analysis Standards and Pipelines Hub can advise and give support. Guidance: Communicating quality, uncertainty and change
Errors and problems are hard to find and fix when they happen	Complex manual processes are usually hard to quality assure. If you find that it is hard to find where issues occur, and takes a lot of time to fix them, your process is carrying quality risks and could also be inefficient.	The Analysis Standards and Pipelines Hub can advise on transforming and automating manual processes using code. Guidance: Quality assurance of code for analysis and research
It is hard for a new starter to understand the process and pick it up	Analysis that is not well documented is hard to understand and hard to reproduce. If it is difficult for members of the team to understand how your analysis works, what steps it involves and any issues, limitations and assumptions then it will be hard for them to run the process. They may also miss potential risks or errors.	The Data Quality Hub and the Analysis Standards and Pipelines Hub can advise on documenting and ensuring reproducibility.



Quality Red Flags	Why does this matter?	What help is available here?
It is difficult to track who made changes to code or datasets and when and why those changes were made	Good version control ensures that you have a full understanding of when, why and how changes were made to your analysis process. If it is hard to track changes, this makes it hard to retrace steps if there is a problem and means you do not fully understand the process.	Guidance: Quality assurance of code for analysis and research
All or part of the analysis is reliant on a single person	Single points of failure carry significant business risk. If there is only one person who understands how to carry out all or part of the analysis then the process is extremely vulnerable.	The Quality Champion in your area can advise on how to flag this risk in your team or division.
Understanding of how and why the process works as it does is reliant on the knowledge or skills of a single person	Single points of failure carry significant business risk. If there is only one person who understands why the analysis works the way it does then the process is extremely vulnerable.	The Quality Champion in your area can advise on how to flag this risk in your team or division.
The analysis I do contains lots of manual steps, like copying and pasting data from one file to another or moving data between software packages with separate steps in each	Manual processes are often inefficient and are prone to user error - a manual process is inherently more risky than a well-designed automated one. Where these exist, you should recognise the need for extra quality assurance to verify that the results are as intended.	<p>The Analysis Standards and Pipelines Hub can advise on transforming and automating manual processes using code.</p> <p>Guidance: Quality assurance of code for analysis and research</p>
I use one or more of the following systems to produce my analysis: CSDB, SAS, SPSS, Stata, Ingres BAW, Excel, or other legacy systems, and there is no plan to move away from legacy tools	Most legacy systems do not support reproducible analysis best practices, and proprietary tools (such as Excel or Stata) are closed source, meaning we cannot fully understand how they work. Workflows that use legacy and/or proprietary tools carry quality risks.	<p>The Analysis Standards and Pipelines Hub can advise on transforming and automating manual processes using code.</p> <p>Guidance: Quality assurance of code for analysis and research</p>

