Office for
Statistics Regulation

Reproducible Analytical Pipelines

Overcoming barriers to adoption

March 2021

Contents

The role of the Office for Statistics Regulation

As an independent UK-wide regulator of government statistics, we are in a unique position to take a broader look at issues of importance to society and to make the case for improved statistics across organisation and Government boundaries. This is supported by our ability to convene, influence and highlight best practice from other sectors.

This review forms part of our programme of systemic reviews which, underpinned by the Code of Practice for Statistics, are aimed at driving improvements in the public value provided by official statistics.

We want to ensure that statistics provide a robust evidence base for national and local policy development and decision making. We champion the need for statistics to support a much wider range of uses, including by charities, community groups and individuals. They should allow individuals and organisations to reach informed decisions, answer important questions, make the case for change and hold government to account.

Executive Summary

Introduction to the review

Official statistics produced by governments should uphold the highest standards of trustworthiness, quality and value in order to serve the public good. In 2017 we championed the Reproducible Analytical Pipeline (RAP), a new way of producing official statistics developed by the Department for Culture, Media and Sport and the Government Digital Service. This approach involved using programming languages to automate manual processes, version control software to robustly manage code and code storage platforms to collaborate, facilitate peer review and publish analysis.

Since then, we have seen some excellent examples of RAP principles being applied across the Government Statistical Service (GSS), the cross-government network of all those who work on official statistics. However, through our regulatory work we have seen that there are often common barriers for teams and organisations wishing to implement RAP. These include access to the right tools and training and statisticians having the time and support to carry out development work.

In Summer 2020 we set out our intention to further advocate for RAP principles in government statistics as part of our Automation and Data programme. We consider that RAP principles support all three pillars of the Code of Practice for Statistics: trustworthiness, quality and value. Trustworthiness, by increasing transparency; quality, by reducing the risk of manual errors; and value, by enabling analytical time to be spent adding value for users rather than on menial, repetitive tasks. It is for all of these reasons that we are so passionate about promoting the use of RAP.

In Autumn 2020 we launched this review. Our aim was to explore the current use of RAP principles across the GSS, identify what enables successful implementation and to understand what prevents statistics producers implementing RAP. We spoke to a variety of organisations that produce official statistics. This included the Office for National Statistics, UK government departments, devolved administrations, arms-length-bodies and voluntary adopters of the Code of Practice for Statistics. We also engaged with users of official statistics and stakeholders with a supportive or leadership role in this area, such as the GSS Best Practice and Impact Team and the office of the National Statistician. Finally, we drew on other available sources of evidence. These included Civil Service and GSS surveys and findings from our previous regulatory work. More information about how we carried out the review is provided in Annex 1: Approach to the review.

Summary of findings

We found that RAP principles support the highest standards of trustworthiness, quality and value. Previous justifications for implementing RAP have often focused on the cost benefits that result from more efficient processes. In this report we outline the additional benefits which we consider to enhance compliance with the Code of Practice for Statistics. These include opening up work to collaboration and challenge, supporting statistics producers to be appropriately skilled and making best

Office for Statistics Regulation

use of these skills, and reducing risk through robust quality management. We also found that RAP increases opportunities for data sharing and encourages innovation.

There is a huge potential to apply RAP principles both in and beyond official statistics. There are around 800 National Statistics and many more official statistics. Many of these statistics are published regularly, for example on a quarterly, monthly or weekly basis. In addition to official statistics, governments produce a considerable amount of other types of analyses, for example management information or responses to ad-hoc requests. To date, only a small proportion of these statistics are being produced using RAP principles.

In this report we explore the common themes we discovered when statistics producers implement, or try to implement, RAP principles. We found that enthusiastic, skilled individuals pushing work forward and support from the RAP champions network have driven the adoption of RAP principles across the GSS so far. Whilst there are many organisations which aim to implement RAP in some form, whether as a test case to demonstrate the need for RAP or as part of a broader ambition to use RAP across multiple teams and statistics, the pace of change has been slow. We found a handful of organisations where some RAP principles have been implemented relatively widely, for example using programming languages for several statistics publications. In others, all the principles of RAP have been implemented for a single publication but not rolled out more widely. Crucially, we found no organisations who follow RAP principles as their default approach to all analytical work and there are several organisations who are still early in their RAP journey.

We found that there is a lack of clear direction in the GSS to implement RAP more widely. Despite the inclusion of RAP principles in good practice guidance and the sharing of several RAP success stories, there is not a specific strategy focused on rolling out the use of RAP in official statistics. As RAP principles are not the default approach to analytical work in government, most new work is currently delivered using traditional approaches. These can be time consuming, prone to human error and difficult to reproduce and share with others. We recognise that RAP developments require some level of resource to implement and that pressures on statistics producers are high. These pressures have increased in many cases during the coronavirus pandemic. However, organisations that commit to a RAP approach now are investing in significant improvements for the future.

While our focus as a regulator is on official statistics, what we have found in this report applies to any type of government analysis. We recommend that those responsible for other types of analysis should also consider the benefits that a RAP approach would bring to improve the trustworthiness, quality and value of their work.

Office for Statistics Regulation

Summary of recommendations

To enhance the trustworthiness, quality and value of official statistics through increased use of RAP principles and see RAP become the default approach to statistics we make the following recommendations.

No.	Finding	Recommendation
1	A consistent shared understanding of RAP and RAP principles is needed across the GSS.	Building on their previous work to promote RAP, the Best Practice and Impact Team and RAP champions network should ensure that there is widespread awareness within the GSS of the recently developed minimum standard of RAP.
2	RAP is not only a change in tools – it involves a cultural change to the way that analysis is approached and carried out.	The Analysis Function board and Directors of Analysis should consider how best to foster a culture where reproducible analysis is prioritised across government.
3	RAP principles support the highest standards of trustworthiness, quality and value and should be used as a way to enhance compliance with the Code of Practice for Statistics.	The leadership of the GSS, including the National Statistician, should set a strategic direction for the use of RAP principles in official statistics.
4	Support and encouragement from senior leaders allows statistics producers to successfully and sustainably implement RAP.	Organisations in the GSS should ensure that RAP principles are included in their analytical strategies. Senior leaders responsible for strategies in their organisations must develop a good understanding of what RAP is, why it is required, and support an open culture of innovation.
5	The implementation of RAP principles is most successful when producers carry out their own development work and when a planned approach is taken – for example having a good understanding of skill levels, training needs and existing processes.	Statistics producers should take a managed approach to implementing RAP. Projects should be underpinned by senior support, sufficient resource and the required skills, training and mentoring support.

Office for Statistics Regulation

No.	Finding	Recommendation
6	RAP is not all or nothing: implementing just some RAP principles will result in improvements.	Statistics producers should consider what can be achieved easily and build on developments iteratively over time.
7	Programming and code management skills are essential for modern statistical analysis.	The GSS People Committee should ensure that RAP-related skills such as coding and code management are considered core skills for statistics producers and included in future career frameworks, such as the competency framework.
8	Bespoke and targeted training is most successful. Statistics producers need access to advanced training on programming, as well as introductory courses.	The GSS should invest in advanced and bespoke training on RAP and RAP-related skills through the Analytical Learning Team. This should build on existing resources and be developed in collaboration with the Best Practice and Impact Team. Availability of training must be effectively communicated across the GSS so everyone is aware of it.
9	Support from experts has a significant impact on the success of RAP projects.	The GSS needs to invest in expert mentoring, for example, through the Best Practice and Impact Team. Organisations that have the required skills and knowledge should support those that don't.
10	Access to the tools required for RAP, such as programming languages, version control software and code storage platforms, varies across organisations. Organisations are tackling the same technical problems with different results.	A strategy for implementing RAP principles across the GSS should recommend tools which should be available to statistics producers. It should also provide guidance on the best approaches to solving common technical problems.

Next steps

We will continue to engage closely with the groups identified in our recommendations and will report on the progress made towards meeting these later this year. We will continue to advocate for the use of RAP principles in official statistics through both our Automation and Data programme and our regulatory work.

Office for Statistics Regulation

To support this, we have developed internal guidance on RAP principles. We have shared this guidance in Annex 2: Regulatory guidance on RAP principles, setting out the kind of questions statistics producers can expect from us on this topic in future. This guidance may also be a helpful self-assessment tool for statistics producers to guide their thinking about the approach they take to their work.

Chapter 1: The need for RAP in official statistics

Official statistics and the statistical system

Official statistics are statistics which are produced by crown bodies, those acting on behalf of crown bodies, or those specified in statutory orders. This includes UK government departments, devolved administrations, arms-length-bodies and public health organisations. Official statistics are a vital part of society, informing public and political debates and touching on topics that include health, education, the economy, crime and the environment. There are over 800 National Statistics – those official statistics that are assessed as fully compliant with the Code of Practice for Statistics – and many more official statistics. Most of these statistics are published on a regular basis, for example quarterly, monthly, weekly or even daily. In addition to official statistics, governments across the UK produce a considerable amount of other types of analyses, for example management information or responses to ad-hoc requests such as Parliamentary Questions or Freedom of Information requests.

The Government Statistical Service (GSS) is a cross-government network consisting of those who work on official statistics. The GSS comprises statisticians, data scientists, researchers and others, and is led by the National Statistician. It is supported by teams in the Best Practice and Impact Team within the Office for National Statistics (ONS) which set guidance and offer consultancy services. The network is also supported by the Analytical Learning Team which develops and provides training. Several committees and groups help to govern the GSS, including the Heads of Profession Group which advises the National Statistician and the People Committee which focuses on recruitment, personal development and talent management.

The Analysis Function is a cross-government network consisting of those involved in the generation of all analysis across government, including statisticians, data scientists, economists, operational researchers, analysts, actuaries and many more. The Analysis Function aims to improve the analytical capability of the Civil Service and enable policy makers to easily access advice, analysis, research and evidence, using consistent, professional standards. It is led by the National Statistician and a board comprising the heads of professions.

Reproducible Analytical Pipelines

The Reproducible Analytical Pipeline, also referred to as RAP, is a set of principles and good practices for data analysis and presentation. RAP was developed by statistics producers in the Department for Culture, Media and Sport and the Government Digital Service in 2017[1] as a solution to overcome several problems: in particular, time-consuming and error-prone manual processes, and an overreliance on spreadsheets and proprietary software for data storage, analysis and presentation. RAP combines modern statistical tools with software development

---

[1] Blog on Transforming the process of producing official statistics

Office for Statistics Regulation

good practice to carry out all the steps of statistical production, from input data to the final output, in a high quality, sustainable and transparent way.

A minimum standard of RAP has recently been developed by the Best Practice and Impact Team. To meet this minimum standard the following principles are required:

- **Peer review** to ensure the process is reproducible and identify improvements

- **No or minimal manual interference**, for example copy-paste, point-click or drag-drop steps – instead the process should be carried out using computer code which can be inspected by others

- **Open-source programming languages**, such as R or Python, for coding so that processes do not rely on proprietary software licenses and can be reproduced by statistics producers and users

- **Version control software,** such as Git, to guarantee an audit trail of changes made to code

- **Publication of code**, whenever possible, on code hosting platforms such as GitHub to improve transparency

- **Well-commented code** and **embedded documentation** to ensure the process can be understood and used by others

- **Embedding of existing quality assurance practices in code**, following guidance set by organisations and the GSS

These fundamental principles can be enhanced, for example by writing code in modular functions that can be reused, or introducing unit tests to ensure that code works as expected. It is also important to note that RAP is not all or nothing – implementing just some of the principles above will generate improvements.

The progress so far

Following the development of the first RAP in 2017, there have been several examples of RAP projects in GSS organisations – for example at the Department for Education, the Centre for Crime and Justice at ONS and the Ministry of Housing Communities and Local Government. Following our blog in 2017, which welcomed RAP as an exciting new innovation, we highlighted RAP developments in a case study under the innovation and improvement principle of the Code of Practice for Statistics.

In 2018 a RAP champions network was established, comprising volunteers from GSS organisations who were enthusiastic about applying RAP principles to their own work and promoting them more widely within their organisations. The network meets a few times a year to share developments and discuss challenges, and is supported by the Best Practice and Impact Team and a Steering Group.

Despite the benefits of RAP being demonstrated by projects such as those above and the acknowledgment of RAP as an important innovation, including at the annual GSS awards, the adoption of RAP principles has been slow. Only a small proportion

10

of official statistics are currently produced using RAP principles. In Autumn 2020, the Best Practice and Impact Team carried out the [Coding in Analysis and Research Survey](#). Out of 1060 responses from across government, nearly 400 people had not heard of RAP and nearly 300 rarely or never use coding in their work. While this is a gap, it also presents a huge opportunity to apply RAP principles and harness their benefits in and beyond official statistics.

The benefits of RAP

Our review has identified a clear need for RAP in official statistics, as RAP principles support the highest standards of trustworthiness, quality and value. Previous justifications for implementing RAP have often focused on the [cost benefits](#) that result from more efficient processes and a reduced number of proprietary software licenses. We have identified additional benefits which enhance compliance with the Code of Practice for Statistics. RAP opens work up to collaboration and challenge, supports the people producing statistics to be appropriately skilled and make best use of their skills, and ensures robust quality management. We consider that these benefits provide a compelling case for following RAP principles when producing official statistics. We explore these benefits in more detail in the following sections.

Trustworthiness

RAP increases transparency and trust in statistics

Confidence in the people and organisations that produce official statistics is vital. Trust in official statistics is supported by opening up internal processes to others and inviting scrutiny via independent measures such as internal and external peer review. All organisations that produce official statistics should seek and exploit opportunities to collaborate, so that knowledge is shared and resources are used most effectively. During the coronavirus pandemic we have seen that [transparency is an essential aspect of public trust](#) in governments and the decisions they make which impact everyone in society. While we have seen a commitment to transparency from statistics producers during the pandemic, we consider there is more that governments can do to increase trust and transparency in statistics and data.

We found that enhanced trustworthiness is the least well-understood aspect of RAP across the GSS, but we consider it to be central to the benefits that are realised by following RAP principles. RAP increases transparency both within organisations and with users outside of organisations as, whenever possible, RAP requires that code for analysis should be published. Traditional ways of sharing information about methodology usually rely on an accompanying methodology report. This can be useful, particularly for a more general user. However, well-written code is a way of communicating with humans as well as computers. Publishing code can be the most effective way to explain the approach taken to expert users. It opens up underlying methodologies and processes, demonstrates what was done and enables external scrutiny.

Public code hosting platforms such as GitHub can be used not just to publish the final version of code, but to host the working version so that the work of statistics producers is carried out in the open. Features such as the Issues log on GitHub

Office for Statistics Regulation

mean that producers can be open about areas they have identified for improvement, as well as inviting ideas for improvements from others. Where it is not possible to make code public, for example in the case of sensitive or confidential content, internally-hosted platforms such as GitLab can be used to share code within an organisation. This is in contrast to the traditional approach where code is stored on local or network drives and usually means that producers cannot see and learn from each other's work outside of their immediate team.

We heard that RAP results in increased collaboration compared to traditional ways of working because peer review is an essential requirement. The purpose of peer review is to ensure that the good practices of RAP have been implemented properly and to provide the opportunity for improvements to be identified. Code hosting platforms provide a transparent and extremely efficient way to facilitate this collaboration. This is in contrast to traditional approaches, which might involve emailing copies of code back and forth between colleagues, or no external scrutiny at all.

Publishing code creates the opportunity for others to reuse it and therefore reduce duplication of effort. As a specific example, responding to the significant demand for new analyses during the coronavirus pandemic would have been greatly supported by widespread publishing of code. This code would have been a valuable resource which could have been shared among producers to save time and create greater consistency, as well as enhancing transparency about methods with the public. Published code also provides examples for other statistics producers to learn and draw inspiration from.

> **CASE STUDY: Taking an open approach to code at Public Health Scotland**
>
> Public Health Scotland (PHS) has published the code for many of its statistics on its GitHub account. While not all PHS analysis can be made public due to sensitive or confidential restrictions, the Head of Profession for Statistics has given strong support to statisticians to take an open approach to their code whenever possible. PHS considers that the risks of sharing code can be mitigated and are outweighed by the benefits of increased collaboration and transparency.
>
> As its first RAP project, Public Health Scotland created an R package to produce its quarterly Hospital Standardised Mortality Ratio statistics. The code for the R package is published on GitHub, allowing expert users to fully understand the underlying methodology. The team responsible for the publication works in the open to produce and develop the publication. The Issues log feature allows users to see what changes are being planned or even to suggest changes of their own. Making the code public has resulted in feedback from others which has led to improvements to the statistics and has facilitated its reuse by colleagues, who use it as a gold-standard reference for their own code.

RAP ensures statistics producers have and use appropriate skills

Confidence in the people and organisations that produce official statistics comes from ensuring that those producing statistics are appropriately skilled, trained and

supported in their roles and professional development. Consideration should be given to future skill needs and staff should be provided with the time and resources required to develop their knowledge and competencies.

By removing manual processes, RAP allows statistics producers to focus on the work they are trained for and specialised in rather than spending time carrying out menial, repetitive tasks. The Coding in Analysis and Research Survey found that nearly half of respondents who can code learnt this skill in education. Given the trend in tools used in statistics and other STEM university courses[2], the use of programming languages and version control software meets the demand from statistics producers who have recently graduated and want to use their existing skills. The adoption of RAP principles also provides the opportunity for statistics producers to learn new skills that they value because they consider them to be skills which will be required in future. We found that, in the teams where RAP has been introduced, statistics producers are more challenged, enjoy their work and feel good about their jobs. As one producer told us, implementing RAP in their team had increased morale because staff felt they had learned the "skills for tomorrow today".

There are additional benefits which can be realised as a result of the professional development of RAP skills. Programming languages are powerful tools for carrying out statistical analysis, but they can also be used to create innovative and user-friendly statistical outputs such as data visualisations. Since many RAP principles derive from software development good practices, we heard that statistics producers also develop the ability to work more closely with software development and web teams on the dissemination of their statistics.

Quality

RAP assures quality through rigorous quality management

Statistics and data must be accurate and reliable in order to meet the needs of users. This requires statistics producers to assure themselves and others that the statistics meet an acceptable level of quality. We expect producers to consider a broad approach to quality assurance, including peer reviews and carrying out deep dives of analyses. While producers of statistics will never eliminate errors entirely, they should strive to establish and maintain effective quality management which reduces the risk of errors.

The Government Data Quality Hub recommends that official statistics should be made reproducible through the use of RAP to make it easier for others to quality assure, assess, critique and re-use methods and results. This recommendation is underpinned by one of the four goals of the GSS Quality Strategy: to implement automated processes to make analysis reproducible.

We found that RAP significantly improves the quality of official statistics by introducing more-effective quality management. It sets a common standard for

---

[2] For example: University of Glasgow Undergraduate in Statistics; University College London Undergraduates in Mathermatics and Statistics; Newcastle University Undergraduates in Mathematics and Statistics; Durham University Undergraduate in Mathematics

Office for Statistics Regulation

analysis and ensures work is sustainable, for example when there are staff changes. Version control software facilitates more rigorous management of the quality assurance process. Statistics producers and their managers are guaranteed an audit trail so they can see what changes have been made, when and by who, as well as who has carried out which quality assurance checks. Version control software ensures that producers save a comprehensive archive of analysis and are able to revert back to previous versions if required. Using traditional ways of working, previous versions of files might have been deleted forever or changes could be made by one person without anyone else knowing.

RAP requires that all the steps in the production process are written in code, rather than carried out manually. This results in a more consistent and logical approach to quality assurance as opposed to using subjective judgment. The use of code also facilitates automated and standardised checking of input and output data. This can be used to identify unexpected changes and stop the process at this point rather than allowing an error to run until it is identified at the end.

Since RAP often requires analysis to be rewritten (for example from Excel, SAS or SPSS into R or Python) we heard that it can improve understanding of existing processes and legacy analysis which may previously have been a black box or only understood by one person. This offers the opportunity to identify and fix errors in the existing process. The good practices required by RAP such as code comments, documentation and an audit trail ensure that statistics are sustainable and robust to changes in staff which is common in many GSS organisations. As discussed in the section on Trustworthiness, the publication of code facilitates peer review and increases collaboration resulting in further improvements to quality.

CASE STUDY: Strengthening quality management at the Department for Transport

To introduce RAP principles to its official statistics, the Department for Transport (DfT) has focused on automating data tables and quality assurance processes. DfT identified these as the best areas for development in its existing processes since they would be the most prone to human error. For example, by using R code to automatically run validation checks and identify issues for further exploration, quality assurance is now carried out in a more standardised and efficient way than it was before for DfT's Road Safety statistics.

DfT ensures that the R code to produce these statistics is peer reviewed, providing an additional layer of quality assurance. Peer review is often carried out by members of the RAP committee, the group which supports RAP developments in the department. The committee has developed a template which is used as the basis for all new coding projects. This supports a standardised coding style across the department and results in improved quality, readability and reusability of code. Finally, DfT has developed a RAP training session for managers which focusses on quality assurance and gives managers the confidence they need to sign off publications which use a RAP approach.

Value

RAP facilitates reuse and accessibility of data and statistics for technical users

Official statistics are an essential public asset and, as such, must be useful and easy to access. This means that statistics and data, as well as supporting information such as metadata and code, should be published at a sufficient level of detail to meet the needs of a range of users. Statistics should be published in forms that enable and enhance their reuse and provide opportunities for data sharing and data linkage.

Our guide to Unlocking the value of data through onward sharing recommends that statistics should be provided in a format which supports data sharing. More-technical users need access to statistics for their own analysis and can find that there is a considerable burden on them to extract and tidy these data from highly formatted Excel tables. This burden increases when the user requires statistics published by different organisations. Examples of these types of challenging data formats include: wide tables where variables are stored in both rows and columns; merged column headers and unhelpful padding; hidden rows and columns; changes in the structure of tables between publications; and time series which are spread across several files. We heard about the impact of this on users in the National Audit Office who currently spend a considerable amount of time converting formatted Excel tables published by governments into flat files which can be analysed, combined and used in their own applications.

Technical users require access to data in a simple, unformatted structure and stored in a machine-accessible format such as a csv, or on an open data platform. This is also one of the requirements for a RAP process, since data must be in a machine-readable format to be analysed and presented using code. Consequently, the data used by producers to populate statistical outputs such as reports and data tables can be easily published as a by-product of a RAP process. An example of this is the unformatted Hospital Standardised Mortality Ratio data published as a by-product of a RAP process implemented by Public Health Scotland.

Another result of producing reusable data as a by-product of RAP is that the data can be included in an integrated data platform. For example, the Connected Open Government Statistics project (now part of the Integrated Data Programme – Dissemination work strand) brings together data on similar topics so that users don't need to navigate different websites to find what they need. The machine-readable, unformatted data makes it easier for technical users to find and use statistics: search engines can look inside and scrape the datasets, and Application Programming Interfaces (APIs) can be created to automatically extract data for innovative services and applications.

As discussed earlier, we consider that publishing the code used to produce statistics enhances accessibility and reuse by others, including other official statistics producers who can reuse and learn from the code.

RAP enables innovation and improvement in official statistics

Statistics producers should demonstrate creativity and innovation to improve their statistics and the insights they offer for users. Keeping up to date with and harnessing technological advances improves the quality and dissemination of official statistics. Collaborating with other producers helps ensure that innovative ideas and good practices are shared.

RAP saves a huge amount of time compared to traditional approaches, particularly those existing processes that include many manual steps. We heard that a RAP approach makes it easier and quicker to create new analyses from scratch. The coronavirus pandemic has created an unexpected and large demand for new timely data and analyses. We heard that producers that were already skilled and experienced in applying RAP principles felt better able to meet the demand for new coronavirus-related statistics.

The considerable time savings that result from a RAP approach mean that analytical time and skill can be better used to increase value for users. This added value could be achieved by spending more time on quality assurance that requires human input and expertise, adding narrative and context alongside figures to enhance insights for users, or producing new statistics and analysis to meet a user need. Furthermore, the skills required to implement RAP principles can be used to create innovative and user-friendly outputs such as data visualisations or dashboards.

CASE STUDY: Using RAP to better meet user needs for ONS crime statistics

The Nature of Crime data tables produced by the Centre for Crime and Justice (CCJ) at ONS previously relied heavily on Excel and SPSS. To reduce manual effort, save time and improve reproducibility, the CCJ replaced the existing process with R and Python code and introduced Git for version control[3].

Implementing RAP principles resulted in a significant reduction in the time taken to produce the statistics: what was originally three weeks' worth of work for thirteen analysts was reduced to under an hours' work for one. The CCJ are also able to create new analysis more quickly (as an example, it took an hour to add nine new tables to the Python pipeline). With the time saved, the CCJ has focused on providing more value for users. This includes publishing a historic time series, adding more measures and granularity to the tables, and developing its survey processes to provide new crime estimates about COVID-19. The team has adapted the code for this project in order to automate the production of other statistics, such as those on violent crime. Overall, implementing RAP has allowed the CCJ to continue to meet its existing output commitments whilst freeing up resources to focus on meeting user needs.

---

[3] The code for the crime tables is available on GitHub and the team has blogged about its RAP transformation.

Office for Statistics Regulation

Chapter 2: Barriers and enablers

In this chapter we share success stories where producers have overcome barriers to implement RAP so that others can learn from them.

We found that producers of statistics encounter many of the same barriers when implementing RAP, though they often tackle them separately within their own organisations. In addition, we identified common enablers which support successful implementation of some or all RAP principles.

The themes we identified broadly reflect three parts of a producers' RAP journey: making the decision to implement RAP in the first place; planning what approach to take for a RAP project; and the practicalities involved in achieving this. We have made recommendations under each theme which we consider will support the sustainable, widespread adoption of RAP principles across the GSS.

Deciding whether to implement RAP principles

The first barrier that statistics producers must overcome is that of deciding to adopt RAP principles. In order for this decision to be made, statistics producers, their managers and leaders within the organisation need a common understanding of what RAP involves and a motivation to implement it. Organisations must foster a culture that supports innovation and invites challenge to the way things are done.

A shared understanding of RAP is needed across the GSS

Until recently there has not been an agreed GSS-wide definition of what comprises RAP. As a result, we found that there is often a misunderstanding about what RAP is, both within and across organisations. This is fundamental: if people don't know what RAP is then they can't implement it.

A common mischaracterisation of RAP is that it is simply automation of a process using a programming language such as R. This misconception ignores the essential good practices which ensure that the automated process is high quality, transparent and sustainable – for example peer review and effective code management using version control software. We found that the value of publishing code is the least well understood principle of RAP and, as a result, is infrequently done or even considered. While these principles often require a culture change, and can therefore be more challenging to implement, they also bring the greatest benefits.

We found that some statistics producers feel that RAP is not for them because they consider it to be relevant only for data science and not official statistics, or because they think it is only beneficial for repeated analysis. The following responses to the Coding in Analysis and Research Survey illustrate the problem: "My role does not include any RAP, that is handled by other data scientists in the team" and "I've only done one-off analyses rather than RAP worthy work so far".

The GSS needs to challenge the perceptions that RAP principles are only for data scientists or only applicable to work which is repeated. The underlying principles of transparency, quality and reproducibility are relevant to all analysis. Even for a one-off piece of analysis, it is good practice to implement version control and peer review,

Office for Statistics Regulation

and to ensure that the work can be fully inspected and reproduced if necessary in future.

To tackle these misconceptions and create a shared understanding of what RAP is and what the benefits are, a consistent and clear definition of RAP and its elements needs to be adopted. Effective communication to producers and senior leaders across the GSS of the [minimum standard](#) recently developed by the Best Practice and Impact Team and the RAP champions network would help to raise awareness of RAP and its importance for official statistics. In particular, the Best Practice and Impact Team and RAP champions network should reach out to those people and organisations not yet using RAP principles. Guidance on how the fundamental principles of RAP can be further enhanced would also help statistics producers to understand which additional practices might be appropriate for their work.

Recommendation 1: Building on their previous work to promote RAP, the Best Practice and Impact Team and RAP champions network should ensure that there is widespread awareness within the GSS of the recently developed minimum standard of RAP.

RAP often requires a change in culture

RAP is not only a change in tools – it is a cultural change to the way that statistics are produced. We found that organisations that were most successful when implementing RAP demonstrated a culture where innovation and professional development are highly valued. We heard that statistics producers can struggle to implement RAP in organisations where there is not an appetite for change and where senior and strategic support is lacking. Our recent [statistical leadership review](#) also underscored the importance of organisational culture. We found that a culture which supports and empowers statistics producers can overcome many structural and infrastructure-related barriers, and that the culture of an organisation is supported by the most senior leaders. Statistics producers need to feel that they can ask why things are done in a certain way and whether it could be better.

We found that in organisations where the implementation of RAP principles has been most sustainable and widespread, three important elements are evident:

- proactive senior support for change

- enthusiastic individuals

- an organisational strategy that includes RAP principles

We found that bottom-up change was often a key driver for initial RAP developments within organisations, but it was not enough on its own to embed widespread change. When a strategic approach to RAP was taken within organisations, it made clear to staff at all levels what the ambition of the organisation was and gave producers the encouragement and backing they needed to make changes to their statistics.

While there has been some strategic support for RAP at a GSS level, for example through the [Quality Strategy](#), there has not been a strong steer from the top of the GSS. To complement a strategic approach within organisations, we consider that the

GSS leadership, including the National Statistician, should provide clear direction to statistics producers to increase the use of RAP principles. The prioritisation of reproducible analysis across government would also help to support this strategic direction, for example in operational analysis, research and economics as well as statistics.

> **CASE STUDY: Creating a culture that supports RAP developments at the Department for Transport**
>
> The Department for Transport (DfT) has fostered a culture of innovation and improvement which has supported the application of many of the core RAP principles to its statistics. This culture has been created through a combination of factors including enthusiastic and driven individuals, strong senior support and strategic direction.
>
> DfT has a strong community of statisticians and its RAP committee has been instrumental in supporting RAP developments. This includes running internal code clubs, inviting external speakers to share learning, and developing training and tools such as an R project template and an R cookbook which provides comprehensive coding examples. Senior leaders, including the Head of Profession for Statistics and managers responsible for teams of statisticians, have a good understanding of the benefits of RAP. As a result, staff are strongly supported to take time to develop new skills and improve their statistics. DfT's RAP developments have been underpinned by a strategic goal to produce most of its statistics using a RAP approach. This has been recognised by the wider department – for example, the RAP committee won the Excellence in Learning award at the DfT 2020 Staff Celebratory Event.

Recommendation 2: The Analysis Function board and Directors of Analysis should consider how best to foster a culture where reproducible analysis is prioritised across government.

Recommendation 3: The leadership of the GSS, including the National Statistician, should set a strategic direction for the use of RAP principles in official statistics.

Senior leaders must support and encourage RAP implementation

Once organisations and producers are aware of and understand RAP, they need to be motivated to implement it. Currently, statistics producers are almost always required to present strong evidence to their organisations to justify RAP implementation. This involves considerable time (up to two years in some cases we heard) and effort on the part of the producers to put together bespoke business cases that demonstrate efficiency savings to senior leaders. In addition to the producers we spoke to, respondents to the Coding in Analysis and Research Survey reflected the need for senior leaders to drive RAP implementation forwards: "You need to push the DDs [Deputy Directors], G6s and G7s toward the directions you are giving in this survey; they control our time and learning development", and "I have been trying to develop and implement RAPs for my team's work for over a year but

the commitment, priority and resources have not been sustained from senior management". We found that senior leaders who understand the need for RAP are able to support statistics producers and help to unblock barriers they face.

RAP principles should be considered good practice for all analysis in government, including official statistics production. They should be seen as a way to enhance compliance with the Code of Practice for Statistics. We would like to see conversations about RAP framed as 'How can we implement RAP?' rather than 'Why should we implement RAP?'.

Recommendation 4: Organisations in the GSS should ensure that RAP principles are included in their analytical strategies. Senior leaders responsible for strategies in their organisations must develop a good understanding of what RAP is, why it is required, and support an open culture of innovation.

Planning how to implement RAP principles

Once the decision to adopt RAP principles has been made, the approach taken to implement them often differs between organisations. While there is not one right way, we found common themes in what works well – and what doesn't – which can help to guide producers as they plan their approach.

A managed approach is most effective

Individual RAP projects and the wider roll out of RAP principles are most successful when a managed approach is taken. It is important to understand the skills within a team or organisation before embarking on RAP implementation, as well as having a good understanding of existing processes and where the biggest benefits will be realised. We found that in some cases RAP developments cannot be used in practice because the team responsible for the statistics does not have the skills or confidence to run and update the process. This is demoralising for those involved and is a poor use of analytical time.

We found that where RAP has been most successful, a managed approach across the organisation has been used. This is underpinned by senior support and an organisational strategy to introduce RAP, and overseen by a central team or committee which can support others to implement RAP principles in their own statistics. This support might take the form of a buddy system where individuals from the central team are partially embedded within another team such as the approach described below at Public Health Scotland; or a more informal offering of trouble-shooting, peer review and technical advice as required, such as the approach taken at the Department for Transport. The modernisation of statistical outputs can also support a strategic approach to implementing RAP – for example, in the case of the Department for Education, where statistics are required to be produced in a consistent structure and format for its dissemination platform.

CASE STUDY: Using a managed approach to implement RAP principles across Public Health Scotland

Following a successful proof of concept project, Public Health Scotland (PHS) established a central transformation team to manage the rollout of RAP principles

across the organisation. The team collected information on self-assessed levels of coding skills, as well as the resource required to carry out existing processes. This allowed strategic decisions to be made about which statistics to focus on first, for example in teams which already had coding capabilities or where there were the most benefits to be gained by automating processes.

Using a buddy system, the transformation team mentored colleagues to implement RAP principles, including coding, version control and peer review, themselves. Before working with teams, skills assessments were carried out and training was targeted where it was needed. The transformation team has developed several tools, hosted on its Knowledge Base, to ensure consistency across PHS, such as coding style guides, templates and a minimum standard of RAP. PHS has applied RAP principles to nine of its official statistics and the transformation team is working with several more teams to embed these practices widely across the organisation.

Statistics producers should be empowered to develop RAPs themselves

The pressures on statistics producers are high. RAP often needs to be implemented within the current capacity available, rather than by gaining extra resource. In some teams and organisations, the pressures of business-as-usual work mean that producers feel unable to take the time to learn new skills required for RAP and redevelop their existing processes. We also heard of developments that are started but then stall due to a lack of consistent resource, or team members leaving. During 2020, the coronavirus pandemic impacted many producers' ability to start or continue RAP developments. However, the pandemic also demonstrated that the statistical system could make resource available and move it flexibly when required. While RAP requires upfront time to carry out development, the improvements to trustworthiness, quality and value, as well as the time savings realised in the longer term, are substantial.

We found that developments are most sustainable when teams are empowered to carry out the work themselves – as opposed to individuals outside their team carrying out the work and handing it over at the end. Dedicated resource with protected time for RAP projects can be helpful in driving forward changes. As such, senior leaders must ensure that sufficient time is made available and prioritised for development and innovation work. To support effective delivery, our statistical leadership review highlighted the need for senior leaders to raise the profile of the work of statisticians and make the case for additional resource where necessary. Senior leaders should also consider whether RAP developments can be managed within existing resources or whether other work can be stopped or paused, without detriment to users.

CASE STUDY: Providing dedicated resource to allow producers to implement RAP principles in ONS crime statistics

Following a successful proof of concept project to demonstrate efficiency and quality improvements to senior leaders, the Centre for Crime and Justice (CCJ) at

ONS established a team to deliver further RAP developments. With agreement from their line managers, members of staff who were interested dedicated two days a week to the team. Support from the Deputy Director and other senior leaders was essential in protecting this time commitment and prioritising development work among competing priorities. This senior support also meant that analysts felt more able to get involved in the project in the first place.

To support the development work, the Best Practice and Impact Team provided mentoring and training. This helped to embed RAP knowledge and skills within CCJ. Despite some initial apprehension about implementing RAP, the team members are confident in the new skills they have developed and feel proud of their work. The CCJ has now applied this approach to mentoring internally, without the support of Best Practice and Impact Team, and is focusing on skills development across the division.

Recommendation 5: Statistics producers should take a managed approach to implementing RAP. Projects should be underpinned by senior support, sufficient resource and the required skills, training and mentoring support.

RAP can be introduced incrementally and iteratively

As discussed in Chapter 1: The need for RAP in official statistics, RAP is not all or nothing. Implementing just some RAP principles, or applying RAP principles to just one part of a statistical process, will result in improvements. In fact, we found that it is often most effective not to try tackling everything at once and instead to take an iterative approach. For example, producers may wish to initially focus on a subset of RAP principles which can be applied across many statistics before introducing the rest. Alternatively, teams may find it helpful to break down the existing process of their statistics production in order to focus on applying all RAP principles to one part first, and then tackle more-challenging sections later. We also found that some producers prefer to start their RAP journey with brand new projects, while others see most value in improving existing processes.

Whatever approach is taken, it is important for statistics producers and senior leaders to be flexible and adapt as teams and organisations learn what works best for them. We found that statistics producers can be creative in meeting the spirit of the RAP principles, even when barriers prevent them from implementing all aspects as required by the minimum standard. For example, while the Department for Transport works on getting access to version control software, it has implemented milestones for code and ensures that code is peer reviewed by members of its RAP committee. In the absence of version control software, this approach ensures that code is managed to a high standard and sustainable.

CASE STUDY: Taking an Agile project management approach to implementing RAP at the Universities and Colleges Admissions Service

The Universities and Colleges Admissions Service (UCAS) is currently in the process of applying RAP principles to its analytical work. In order to create

automated pipelines and introduce version control software and peer review to its working practices, UCAS is following an Agile project management approach. This means that teams take an iterative approach to their work, learning from continuous feedback throughout the development process, collaborating closely with stakeholders and adapting their approach as necessary. Since RAP is a new way of working at UCAS, this has been particularly helpful as it has allowed the organisation to discover what works best for it and its staff.

UCAS decided to start its RAP transformation by applying RAP principles to new projects first, establishing best practices before applying these to legacy code. It has learned that it is most effective to start small and build on developments, and to make sure that all teams are aligned in their approach, rather than trying to implement RAP in separate pockets across the organisation.

Recommendation 6: Statistics producers should consider what can be achieved easily and build on developments iteratively over time.

Putting RAP principles into practice

The final set of barriers and enablers that we identified are those which prevent or support work in practice. These include training in RAP-related skills, as well as the technical tools which are available to statistics producers.

Coding and code management skills are essential for modern statistical analysis

Open-source programming languages provide powerful tools to carry out data analysis and statistical modelling, and to present results. Programming languages also integrate well with version control software to support good code management practices. In recent years, the trend in tools for statistical analysis has moved away from proprietary software and towards open-source programming languages. This can most clearly be seen in academia and data science – for example, the open-source programming language R was developed by academics to support statistical analysis and is popular among data scientists, along with Python. University statistics courses largely teach programming languages such as R and Python and are starting to introduce version control training. Programming languages are also now being used or introduced in other STEM subjects such as maths, physics and psychology.

A move towards using open-source programming languages to produce official statistics would bring the GSS in line with other industries, make use of the skills and experience of new starts, and develop the skills of those working in government. One respondent to the Coding in Analysis and Research Survey reflected that "I think it is becoming more common to need to know code to effectively and efficiently carry out your job role in statistics. The younger generation coming in will develop this a lot too, to the point where it is essential knowledge".

It is important that the GSS attracts, retains and nurtures top analytical talent by ensuring that the right skills and capabilities are prioritised now and in the future. In

Office for Statistics Regulation

December 2020 the GSS People Committee published a [pilot version of the competency framework](#), with the goal of refreshing the framework for the 2021/22 financial year. As the new framework will be used in future recruitment, performance management and development discussions, we consider it essential that RAP-related skills such as coding and code management are included in the new version.

Recommendation 7: The GSS People Committee should ensure that RAP-related skills such as coding and code management are considered core skills for statistics producers and included in future career frameworks, such as the competency framework.

Training and guidance should be bespoke and targeted

A change in tools requires training to be available for statistics producers who are unfamiliar with them. For many, moving from proprietary software syntax to a programming language is a fairly straightforward change, but others will require more-comprehensive training. Managers responsible for statistics also need to have a sufficient level of knowledge about code to confidently quality assure and sign off publications.

Statistics producers currently have access to introductory courses on programming languages and version control software – for example through the [Analytical Learning Team](#) or other online resources produced outside government. However, we heard that this kind of training is often too generic for producers to apply to their own work or not advanced enough to support producers to use these tools for all their analysis. A strong theme that emerged from the Coding in Analysis and Research Survey was that more training on RAP-related skills, including code management and other software development good practices, is needed. For example, one respondent wrote: "Although there has been some RAP training, it is not enough for me to feel confident to even try it" and another said: "I desperately want to use unit testing, but don't know how". We found that some people and organisations are not aware of existing training offered for the GSS, so it is important that the availability of existing training is more widely communicated. A list of RAP-related training is included in Annex 3: RAP-related training resources.

We found that some producers, such as Public Health Scotland and Department for Education, have created their own bespoke training so that staff can apply what they learn to their work more easily – and in some cases can learn and then implement RAP principles themselves or with minimal support. The Department for Transport has also developed bespoke training for managers on RAP so that they feel confident in quality assuring new processes. Although these resources are hugely beneficial, creating these kind of bespoke training courses takes time which we heard is not always available to producers.

It is important that training is targeted in a strategic way so that statistics producers can put what they have learned into practice. We found that often producers take training courses without a project to focus on afterwards – if the skills learned are not reinforced then they can be quickly lost. The Analytical Learning Team offers a service for producers to help them plan their approach to training and we encourage producers to make use of this.

Office for Statistics Regulation

Finally, developing internal guidance and standards for coding will ensure that statistics producers understand what is expected of them and how to achieve it. Guidance also supports analytical work to be done consistently and sustainably – for example, recommending what coding style or R packages should be used so that work can easily be passed among colleagues.

> **CASE STUDY: Creating bespoke guidance and training at the Department for Education**
>
> At the Department for Education (DfE) bespoke tools and guidance have supported progress in implementing RAP principles, with an initial focus on introducing programming languages and good coding practice. The DfE statistics development team is responsible for supporting staff to develop RAP-related skills and apply RAP to their work. The team discovered that there was a lack of clarity among staff about what RAP is and that many considered it to be only relevant for data science.
>
> To help with understanding and consistency, the statistics development team created a bespoke minimum standard of RAP for DfE. A self-assessment tool based on this minimum standard was developed which other teams use to monitor their progress towards it. This has made RAP feel more achievable and the tool also provides an option for teams to request further training. The statistics development team has produced walkthrough videos on using R at DfE, which helped staff to get started with R and to ensure standards across the department.

Recommendation 8: The GSS should invest in advanced and bespoke training on RAP and RAP-related skills through the Analytical Learning Team. This should build on existing resources and be developed in collaboration with the Best Practice and Impact Team. Availability of training must be effectively communicated across the GSS so everyone is aware of it.

Statistics producers value expert and peer-to-peer support

In addition to training courses, we found that support from experts has a huge impact on the success of RAP projects. This expert support can take many forms, including buddying or mentoring throughout a project, pair programming to write code and offering troubleshooting advice when things go wrong. This type of support is currently provided in-house through central RAP support teams or code clubs in some organisations, or more widely across the GSS by the Best Practice and Impact Team. However, not all organisations have individuals with the required skills to provide this type of support, and, to date, the Best Practice and Impact Team have only been able to support a small number of projects with the resources they have had.

We encourage statistics producers to share knowledge so that individuals, teams and organisations can learn from and inspire each other, as well as provide technical help. One respondent to the Coding in Analysis and Research Survey reflected that "It would be great if existing teams who have the skills and experience could teach across organisations in some way".

Office for Statistics Regulation

The Best Practice and Impact Team oversees the [RAP champions network](#) which consists of volunteers from GSS organisations who meet a few times a year to share new RAP developments. The RAP champions network has been a valuable way to share learning and provide technical peer support. However, out of 1060 responses to the Coding in Analysis and Research Survey, only 179 people had a RAP champion in their organisation and knew who they were. We consider that making the RAP champion role more formal within organisations would help to raise awareness of RAP across the GSS and make better use of the support and knowledge the network can provide – both to their own organisations and the wider GSS. The Government Data Science Slack workspace also hosts a RAP channel where producers can ask each other questions and discuss their work. We encourage all statistics producers who are embarking, or thinking of embarking, on a RAP development to reach out to the Best Practice and Impact Team and the wider RAP network.

Recommendation 9: The GSS needs to invest in expert mentoring, for example through the Best Practice and Impact Team. Organisations that have the required skills and knowledge should support those that don't.

A consistent approach to tools across the GSS would help

Access to the tools required for RAP varies across organisations in the GSS. Though in the past, access to programming languages such as R and Python has been a significant barrier preventing producers from starting their RAP journey, these tools are now more commonly available. However, a lack of access to version control software and code hosting platforms still presents a significant barrier to implementing some of the basic principles of RAP. In particular, access to public code hosting platforms such as GitHub is challenging. A few organisations have permitted the use of GitHub for publishing code and working in the open, such as the Scottish Government and ONS, but many others do not allow access due to security concerns. While publishing code and working in the open does involve a level of risk, there are ways to mitigate this risk so that the benefits of publishing code can be achieved. We found that obtaining access to platforms such as GitHub requires support from, and collaboration with, colleagues outside statistics teams, for example senior leaders and IT departments.

We found that building a strong working relationship with IT colleagues can help statistics producers obtain the technical infrastructure they need. However, we found that producers are often tackling the same technical problems within their own organisations with different results – whether this is agreement on how R packages can be downloaded and updated or access to code hosting platforms or collaboration tools like Slack. A common GSS-wide approach to all these issues would be challenging since each organisation has its own IT department. Instead, a GSS strategy on implementing RAP should include recommendations on tools and the best approaches to solve common problems. This would equip statistics producers with the knowledge and leverage they need to affect change in their own organisations.

CASE STUDY: Using Git and GitHub at the Scottish Government and National Records of Scotland

The Scottish Government and National Records of Scotland use a GitHub organisation to publish some of their analysis. In the last year this has included COVID-19 data and code, such as weekly COVID-19 management information. The GitHub organisation also hosts outputs from RAP projects, such as the Scottish Government Exports Performance Monitor.

The Scottish Government has recently implemented a virtual workbench which enables users to connect to GitHub and a host of other analytical tools from their work laptops. Implementing this solution required senior leaders to understand the issues staff were facing and support changes to the IT infrastructure. Implementation was also accelerated by the demand for new and frequent analysis about the coronavirus pandemic. Scottish Government has found that there has been good uptake in the use of the workbench, and this will help to support its ongoing RAP developments.

Recommendation 10: A strategy for implementing RAP principles across the GSS should recommend tools which should be available to statistics producers. It should also provide guidance on the best approaches to solving common technical problems.

Annex 1: Approach to the review

Engagement

An important part of this review involved talking to statistics producers, those with a supportive or leadership role in the GSS and users of government statistics. We identified people to speak to through various routes, including previous regulatory work and inviting expressions of interest on our website and government Slack workspaces.

Below is a list of the organisations and teams we engaged with during the review. We would like to thank everyone who gave their time to speak with us.

- Cabinet Office (producer)
- Centre for Crime and Justice, ONS (producer)
- Connected Open Government Statistics, ONS (producer)
- Department for Business, Energy & Industrial Strategy (producer)
- Department for Education (producer)
- Department for Infrastructure (producer)
- Department for Transport (producer)
- Department for Work and Pensions (producer)
- Economics Statistics Development, ONS (producer)
- Essex County Council (producer)
- Best Practice and Impact Team, ONS
- Analytical Learning Team, ONS
- GSS People Committee
- Home Office (producer)
- Institute for Government (user)
- Office of Rail & Road (producer)
- Public Health England (producer)
- Public Health Scotland (producer)
- Public Health Wales (producer)
- National Audit Office (user)
- National Records of Scotland (producer)
- Northern Ireland Statistics and Research Agency (producer)
- Scottish Government (producer)
- Universities and Colleges Admissions Service (producer)

Other sources of evidence

In addition to the interviews above, we gathered evidence from other sources. This included the results of surveys (such as the Civil Service People Survey, Coding in Analysis and Research Survey and Statistical Quality in the GSS survey), information published by the RAP champions network and in blogs and written contributions from statistics producers and academics.

Office for Statistics Regulation

Annex 2: Regulatory guidance on RAP principles

This annex includes questions that statistics producers can use to understand their current use of RAP principles and where there may be areas for improvement. These are the kind of questions that statistics producers can expect from us on this topic during regulatory reviews and assessments in future.

Processes and tools

What are the steps in your production process, from acquiring the data to final outputs? Can you map out the process? Why is it done in this way?

How many manual steps are there in the process (e.g. updating cells in spreadsheets, moving data between software or copy-paste steps)? Could these be reduced to minimise the risk of error?

How do you ensure that input data are correct and in the expected format and structure?

What analytical tools do you use during the production process? Are they the best for the job?

Version control

What is your approach to version control? Do you use version control software?

How do you ensure that work carried out by your teams is reproducible? For example, could you reproduce the numbers from a previous publication?

How do you ensure that work carried out by your teams is auditable? For example, if changes need to be made how are these documented? Would you always know who made the changes, when and why?

Peer review and working in the open

How many people are involved in the production process? Are analyses and results checked by another person? Are changes to the analysis checked by another person?

Is your analysis ever peer reviewed by someone outside your team or organisation?

Do you publish the underlying code, either within your organisation or publicly?

Annex 3: RAP-related training resources

This annex includes training resources that cover RAP principles. This list is a snapshot of resources that are currently available. New resources may become available as they are developed.

RAP principles cover many domain areas including computer science and project management, and therefore most courses only focus on particular elements of RAP such as programming or version control. This list is not exhaustive, and many RAP principles may be covered in training courses not specifically aimed at RAP.

Learning Hub

The Analytical Learning Team provides RAP-related training for the GSS. The RAP training has been broadly devised using the BPI [Quality Assurance of Code for Analysis and Research guidance](#). RAP related courses delivered by the Analytical Learning Team can be found on the [Learning Hub](#):

- [Git (Version Control)](#)
- [Best Practices in Programming](#)
- [Unit Testing](#)
- [Packaging and Documentation](#)
- [Introduction to Continuous Integration](#)
- [Awareness of Coding Tools](#)

These have been run in addition to introductory and advanced courses on specific programming languages such as [R](#), [Python](#), and [SQL](#) as well as [command line basics](#).

Massive Open Online Course (MOOC) Providers

Massive Open Online Courses are those that are accessible by anyone online. Some MOOCs comprise traditional training material such as recorded lectures and example problems, and some have interactive courses with forums for participants to chat to one another. Examples of MOOCs are: [edX](#), [Udemy](#) and [Coursera](#). Specific 'RAP' courses are uncommon, but free-to-enrol RAP-related courses are available, including:

- [Reproducible Analytical Pipelines (RAP) using R](#) (from Government Digital Service) on Udemy
- [Automated Software Testing: Unit Testing, Coverage Criteria and Design for Testability](#) on edx
- [Intro to Git](#) on Udemy
- [Introduction to Git and GitHub](#) on Coursera
- [Building R Packages](#) on Coursera
- [Testing and Debugging Python](#) on Coursera

Office for Statistics Regulation

MOOCs also comprise paid courses that may be of higher quality and provide the learner access to more material, a shared learning environment, and a certification.

Other Resources

Single lectures, training slides and material on RAP-related skills include:

- [R Style Guide](#) by Hadley Wickham
- [PEP 8 Style Guide for Python](#)
- [Git handbook](#) by GitHub

Office for Statistics Regulation