University of the West of England
Data Research Access and Governance Network (DRAGoN)

Output checking for practitioners

Module handbook
Version 01

Contents

Course aims, objectives, and audience

Who this course is for

This course is intended for those who have responsibility for checking the outputs of researchers, and who already have active experience. The associated certificate is required for those responsible for output-checking in organisations with authority delegated by ONS.

Attendees are assumed to have basic statistical skills: that is, one should understand

- the difference between frequency and magnitude tables
- basic descriptive statistics: mean, median, mode, percentiles, chi-square
- summary statistics arising from analytical modelling (R^2, estimated standard error, F- and t-statistics etc)
- basic graph types: line, bar, x-y, box-and-whisker,
- what estimation models (regressions, ANOVA) are and how they are presented
- the difference between linear and non-linear models
- what the residuals from models represent

Attendees should have passed

- Safe Researcher Training
- Introduction to Output Checking

or otherwise have a practical familiarity with these statistical disclosure control (SDC) concepts:

- Primary disclosure
- Secondary disclosure
- Dominance
- Class disclosure
- The statbarn classification model
- Principles- and rules-based output SDC

Attendees should have practical experience of clearing outputs for researchers. Attendees should be prepared to come and discuss their experiences and reflect on how they work with researchers.

Objectives

Specific learning objectives for the course are for the attendee to develop the skills to

- use his or her own judgment in assessing outputs
- have confidence in dealing with known and unknown types of output
- understand the subjectivity of the process
- develop a strategy for dealing with difficult outputs or unknown types
- develop skills in dealing with researchers
- develop internal consistency when dealing with variable outputs and researchers
- integrate into and help develop the output checking community

Course structure and joining instructions

The course is broken into two parts

Part I: technical skills	This session will focusing on the statistical skills. Exercises of increasing complexity will test and build understanding of the technical aspects of the clearance process – including what to do when you don't know the answer.
Part II: user management	This session will focus on developing an understanding of the human elements of the clearance process. Key to this will be understanding how this is an operational issue.

The course will take approximately 4 hours. The course will be taught online – the link details will be sent out to you as an invitation.

We will be using slides and virtual whiteboard, so please log in via a laptop or tablet; a mobile phone screen will be too small.

The session will be recorded; a link will be circulated afterwards for you to download and review the session, and may be used by the DRAGoN team for pedagogical purposes. If you have concerns about this please let the course organiser know.

In preparation, this course requires **pre-reading** and **preparation of answers to dummy output requests**. The pre-reading will be tested at the start of the course (see **preparation exercise below)**, and attendance is conditional on this.

Assessment

The course will be followed by an on-line assessment. A pass is necessary for successful completion of the course. In the event of failure, the tutor will contact you to discuss further options. **A previous test with model answers will be supplied to you as pre-course required reading.**

Pre-course exercises

You will be sent a dummy output request, intended to represent a typical output request of varying quality. **Prepare to discuss** this dummy output (this models the test structure):

- Which elements give confidence that there is no practical disclosure risk?
- Which elements give concern that there is a practical disclosure risk?
- Should you approve it, reject it, or ask researcher the more information?
- What feedback will you give to the researcher if not approved?

You may wish to consider:

- Does this breach the rules?
    - If so, does it really matter?
    - If not, is there still some risk that means you reject it?
- Is there a link between outputs?
- Do I need any more information? Will it tell you anything you can't reasonably suppose?
- How have you reacted to similar outputs in the past?
- What type of 'statbarn' is it?
    - If safe, are there any conditions attached?
    - If unsafe, what are the relevant considerations?
    - Are you comfortable about safe/low review definitions?
        - if not, why not (be specific)?
- What happens if you're wrong?

Feel free to discuss your answers with others who have been given the same exercise. Note that the important part of these exercises is **understanding the thought processes** that lead you to the decision, not necessarily the decision itself.

Required pre-course reading

The **minimum** level of knowledge required to attend the course is:

- SACRO (2023) SACRO Guide to Statistical Output Checking. Project report. Bristol, October https://zenodo.org/records/10054629 sections 1, 2, 5, plus awareness of the structure of section 7 (the actual statbarns)
- ONS (2019) "Safe researcher training: SDC, technical" http://www.saferesearchertraining.org/SRT_full_version_v00.12%20b%20sdc%20theory.pptx
- ONS (2019) "Safe researchers training: RBOSDC and PBOSDC" http://www.saferesearchertraining.org/SRT_full_version_v00.11%20e%20pbosdc.pptx

For the two slide packs, **read the notes. This level of knowledge will be assumed**. If you struggle with these, you should consider whether you are ready to attend this course.

Finally, please read Model Answers to the test question supplied to you; this is a previous test given to course attendees. Pay **particular attention to 'what candidates got wrong'**, particularly in respectively of uncovering the 'chain of events' that could lead to a breach.

You may also find these two **helpful** but they are **not required**:

- Alves, K., & Ritchie, F. (2020). Runners, repeaters, strangers and aliens: operationalising efficient output disclosure control. Statistical Journal of the IAOS, September https://uwe-repository.worktribe.com/output/6634925/ - a detailed review of how OSDC should be seen as an operational process
- Green E., Ritchie, F. and White P. (2024) The Statbarn: A New Model for Output Statistical Disclosure Control. Privacy in Statistical Databases 2024. Temporarily available from https://bit.ly/3An2bbI

After the course, additional reading materials to support your work will be circulated.

Contact details

For queries about the course contents, delivery, reading, assessment, contact the lecturers (Elizabeth Green and Felix Ritchie) at dragon@uwe.ac.uk. For queries about course administration, location, dates, times etc, contact your local organiser.