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

Output checking
Introductory course

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. It is also 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 also have passed Safe Researcher Training or otherwise have knowledge of the following statistical disclosure control (SDC) concepts:

- Primary disclosure
- Secondary disclosure
- Dominance
- Class disclosure
- Principles- and rules-based output SDC

If these topics are new, please ensure that you have done the Required Reading before attending.

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
- understand the subjectivity of the process
- develop a strategy for dealing with difficult outputs or unknown types
- develop an awareness of the importance of researcher management

Course structure and joining instructions

This course will focus 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

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. 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.

Pre-course exercises

You will be sent 2 dummy output requests, intended to represent a typical output request of varying quality. Others on your course will also get dummy output, but not all attendees will receive the same ones. **Prepare to discuss** these dummy outputs (treat each one as a separate submission):

- 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?

The first example is designed to test basic concepts. 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?
- Do I need any more information? Will it tell you anything you can't reasonably suppose?
- What are the consequence if you make a mistake?

The second example introduces more complex outputs. For this set outputs consider:

- Is this a 'safe' or 'unsafe' statistic?
    - If safe, are there any conditions attached?
    - If unsafe, what are the relevant considerations?
- What happens if you're wrong?
- Are you comfortable about safe/low review definitions?
    - if not, why not (be specific)?

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 either Safe Researcher Training (provided by SCADR, ONS, UKDA or CDRC), or these two papers

- Lowthian P. and Ritchie F. (2017) "Ensuring the confidentiality of statistical outputs from the ADRN" https://uwe-repository.worktribe.com/output/888435/
- Green E. and Ritchie F. (2021) Statistical disclosure control for HESA: Part 1: Review of SDC theory https://uwe-repository.worktribe.com/output/10621148/ (except sections 3.5, 3.6 and 3.8)

If your memory is rusty about SRT, these are the relevant slides – and please **read the notes**:

- ONS (2019) "Safe researcher training: SDC, technical" http://www.saferesearchertraining.org/SRT_full_version_v00.12%20b%20sdc%20theory.pptx

You should look at the SACRO Guide to familiarise yourself with the contents and structure:

- SACRO (2023) Guide to Statistical Output Checking. https://zenodo.org/records/10054629

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 DRAGoN team at the University of the West of England, Bristol (email dragon@uwe.ac.uk). For queries about course administration, location, dates, times etc, contact your local course organiser.