Research Lab Import and Export Guide

for Approved Users

Updated 20260718 by Technical Manager

Data Imports

JDI: Please contact your data scientist to bring any additional data or scripts into the JDI.

DSH: You can import data/scripts directly yourself into the DSH, using the File Transfer Portal website at https://filetransfer.idhs.ucl.ac.uk/webclient/Login.xhtml - however you must first notify us before you import the data/scripts. To do thi, fill out a GeoDS Import Notification Form which can be found at https://service.geods.ac.uk/import_notification_form.docx and send it to data@geods.ac.uk. You can also contact us at this address if you wish to discuss a proposed import, for example if you are concerned about licence implications or you are concerned that this could create a potential data linkage that was not outlined in your approved project proposal form.

Data Exports

Data or model results may only be output from our secure labs (such as DSH or JDI) following approval by the Research Approvals Group (RAG), Senior Management Team (SMT) and two approved GeoDS Data Scientists (DS1 and DS2). A "User" is a GeoDS Approved User who has had a project approved by RAG and has agreed to the terms of the GeoDS User Agreement. See Appendix A for output data requirements.

The steps below must be followed to remove any data from the lab. *The first 5 steps below will likely have already been completed, if you are ready to output data.*

1. *User describes anticipated outputs in Project Proposal Form.*
2. *General outputs approved by the Research Approvals Group (RAG).*
3. *User agrees to the terms of the GeoDS User Agreement.*
4. *User agrees to the terms and conditions of the secure lab.*
5. *User undertakes analysis in lab.*

6. User prepares output files and saves those in a special folder for the output (e.g. your name or the date) in the **R**:/[YOUR_PROJECT_SHARE]/ folder (for DSH) or a folder named 'Final' in User's account (for JDI) using the convention: [DATE]_[FIRSTNAMESURNAME]_[PROJECT]_[FILE_DESCRIPTOR].
    a. The files output should be tabular data only. Disclosure checks will be applied to all data removed from the secure environment. Scripts/code can be exported, these should be in TXT format files.
    b. Any graphs or graphical output should ideally be done outside of the secure environment.
    c. Exporting spatial data should be avoided if possible. If you MUST export spatial data, it must a human readable (e.g. GeoJSON, GeoJSON Newline Delimited or WKT), and will be examined in Notepad and QGIS by the data scientist.

7. User submits an output request at https://service.geods.ac.uk/outputchecks/ and then emails GeoDS Technical Manager at data@geods.ac.uk to notify that such a request has been submitted.

8. Two GeoDS Data Scientists (DS1, DS2) are assigned (from list of lab approved GeoDS Data Scientists) to carry out a check of output files to ensure there is no risk of disclosure. We aim to carry out this check within 5 working days of notification of the submission of the output. The check is normally carried out during a Teams meeting with you, but may be carried out before this or instead of this. They either *approve*; request *minor*/*major revisions*; or *reject* outputs. Their decision is returned to the Technical Manager and if *minor revision* decision is returned to you. If decision is *major revision*, *reject* or *approve* this is forwarded to the Senior Management Team.

9. Senior Management Team reviews the output request. Requests are considered alongside the RAG approved outputs. SMT will approve; approve with revisions; or reject request. If outputs request differs substantially to that approved by RAG the request will be referred back to RAG for approval. The Technical Manager or Centre Manager can act on behalf of SMT to authorise these outputs. If an external user is exporting large amounts of spatial data, then this must be explicitly authorised by the full SMT.

10. If *revise* Technical Manager will inform user of decision. User can then make amendments and resubmit output request, which the Technical Manager to forward to DS1&2. If *reject* Technical Manager will inform user of decision. User may if authorised carry out further analysis with the data and submit a new output request for review (return to step 6).

11. If *accepted* Technical Manager updates the output request appropriately and forwards to (for DSH) the DS to release the data or (for JDI) to User to take to Lab Manager/DS. Lab Manager/DS releases approved files on encrypted media either to User or to DS. If released to DS then they will arrange the delivery of the files to User by secure download.

12. User signs the declaration form attached to their output request on the https://service.geods.ac.uk/outputchecks/ website acknowledging receipt of data and conditions attached to it. The Technical Manager will then file the form internally. If user wishes to have a completed copy of the form, with the review comments and declarations, then this can be requested at data@geods.ac.uk.

Geographic Data Service

Appendix A: Statistical Disclosure Control - Output Requirements

You should be aware of disclosure control considerations from your Safe Researcher Training. The below is a summary of some guidelines that our data scientists, who have also undergone Safe Researcher Training, may use when deciding whether your proposed output is safe to release.

Outputs

Outputs requested should be 'finished outputs' i.e. the finished statistical analyses that you intend to present to the public. If requiring intermediate outputs for a particular reason e.g. to present initial findings then these may be considered if clearly presented and clearly explained. All outputs must be easy to read and interpret and how they are to be used explained.

Non-Disclosive Data

Taken from GSS/GSR Disclosure Control Guidance for Tables Produced from Surveys, October 2014

Social Surveys

- For the majority of surveys, outputs should be for large geographical areas, e.g. Country or Government Office Region, or in some cases Local Authority District (or equivalent). The level of geography should reflect survey design.
- Suppress or combine unsafe cells, i.e. where there are one or two units contributing to the cell.
- Where the sample size of a total or sub-total is one or two, suppress the whole row or column to which the total refers, including any zero cells (or combine neighbouring categories).
- In unweighted tables, cell suppression does not provide sufficient protection. Unsafe cells should only be combined with other cells.
- If unweighted sample base numbers are essential they should be conventionally rounded to base 10.
- Percentages may be released, provided it is not possible to deduce where only one or two units have contributed to the cell.
- Units may be individuals, families or households, communal establishments or any other unit whose confidentiality should be protected.

Subsamples

- For the majority of surveys, outputs should be for large geographical areas, e.g. regions, or in some cases Local Authority District (or equivalent). The level of geography should reflect survey design.
- Table design should be used to remove all unsafe cells, i.e. where there is one unit contributing to a cell. Variable categories should be combined or variables removed until only safe cells remain.
- Percentages may be released, provided it is not possible to deduce where only one unit has contributed to the cell.
- Units may be individuals, families or households or any other unit whose confidentiality should be protected.

Business Surveys: Magnitude tables

- A cell meeting both the following criteria is safe (otherwise the cell is unsafe):
  - there must be at least *n* enterprise groups in a cell (threshold rule)
  - the total of the cell minus the largest *m* reporting unit(s) must be greater than or equal to p% of the value of the largest reporting unit (p% rule)

Note that the values of the p% and minimum threshold parameter *n* and m should remain confidential, since knowledge of these values reduces the protection. The choice of p, n and m would usually be decided by the Responsible Statistician. Typical examples would be 2,3,4,5 (for n), and 2,3 (for m) and 5% 10% ,15%, 20% (for p).

- Table design should be used first to reduce the number of unsafe cells in a table where this is consistent with the main uses of the data.

- Cell suppression is the standard method used to protect tables with unsafe cells. The unsafe cells are suppressed, known as *primary suppressions*. Other cells must be suppressed to prevent the values of the unsafe cells being calculated by subtraction from the marginal totals of the table. These are known as *secondary suppressions*.
- Cell suppression does not generally provide protection from disclosure by differencing. Tables should be published using fixed categories to avoid disclosure by differencing. For example the same geographies and SIC codes should always be used.

Business Surveys: Count data

- Tables of count data are to be protected by redesign of the table to protect sensitive cells. If further protection is required other techniques such as controlled rounding to base 5 should be considered.
- Percentages or rates must be derived from rounded values.

Geographic Data Service

Appendix B: GeoDS Import Notification Form

Geographic Data Service (GeoDS)

Trusted Research Environments
Import Notification Form

1. Contact Details

1.1 Name of User/Researcher: Click here to enter text.

1.2 Secure Facility (UCL DSH, UCL ARC TRE or UCL JDI): Click here to enter text.

---

2. Project

2.2 Project Ref: Click here to enter text.
2.3 Project title: Click here to enter text.

Geographic Data Service

3. Import

3.1. List of files for import

File Name	Description of file contents	File format	(For JDI only) Checked by GeoDS DS

For JDI imports only, file name must be in format:
[DATE]_[FIRSTNAMESURNAME]_[PROJECT]_[FILE_DESCRIPTOR]

3.2. Brief Reason for import: Click here to enter text.

3.3. Transfer Media (for JDI only):

☐ USB Stick ☐ University filestore ☐ CD/DVD ☐ External disc ☐ Email

3.4. Encryption:

Is Transfer Media encrypted? ☐ Yes ☐ No
If so please provide details: Click here to enter text.
Should file be encrypted? ☐ Yes ☐ No
If so please provide details: Click here to enter text.

4. Authorisation (for JDI only)

4.1. I confirm that the data to be imported is as described as above.

GeoDS Data Scientist/Researcher:

Print:

Signature:

Date: