Department for Education

Statistical Notice

Statistical Policy Statement on Confidentiality

Principle 5 of the Code of Practice for Official Statistics requires all producers of Official Statistics to publish transparent guidance on their arrangements for protecting confidential data. The Code is at:

www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html

Version number	Date	Details
1	February 2010	First edition
2	November 2010	Addition of Annex A and reference to annex. Amended references to Department for Children, Schools and Families, to Department for Education.
3	June 2011	Update Head of Profession Update references to suppression of exam results in Annex A
4	April 2013	Deletion of references to the Research and Statistics Gateway

Version 4: Published April 2013

This statistical policy will be reviewed annually.

Confidentiality

The Department for Education is committed to maintaining the confidentiality of the data it receives, stores, processes and disseminates.

We are committed to protecting confidentiality to:

- Maintain the trust and co-operation of those who own and manage administrative data sources used by us and respondents to our surveys.

- Comply with the relevant legislation, including the Data Protection Act 1998.

- Comply with Principle 5 of the Code of Practice for Official Statistics, which states that: “Private information about individual persons (including bodies corporate) compiled in the production of official statistics is confidential, and should be used for statistical purposes only”.

Our arrangements for protecting confidentiality fall into three areas, as follows:

Personnel

All staff who work with data about individual persons, e.g. pupils in schools, receive appropriate security checks and training in protecting information. Secure working areas are provided for staff who work with confidential data about individual persons.

Data

Our arrangements for protecting private information about individual persons include:

- We provide detailed data security operational guidance for our staff.

- We use secure data transfer methods to transfer data to and from external bodies such as local authorities.

- Contracts with external organisations include data security clauses where appropriate.

- External bodies who wish to access our data are required to complete a confidentiality agreement. Where the agreement is approved by the Head of Profession for Statistics, only the minimum data needed for the specified purpose are released.

- We operate Service Level Agreements when sharing data with internal users and other governmental groups.

- Respondents to our data collection exercises receive privacy notices detailing what the data will be used for and our undertakings with respect to confidentiality.

- Raw data cannot be downloaded from our systems; raw data cannot be downloaded onto CDs, through USB ports, to laptops etc.

Statistical Disclosure Control

We identify three types of disclosure risk in relation to the data about individual persons or the statistics derived from the data:

- **Identity**: If a person or persons can be identified (by either the persons themselves or someone else) then there is an identity disclosure risk.

- **Attribute**: If confidential information about a person or group of persons is revealed and can be attributed to the person or each person in the group then there is an attribute disclosure risk.

- **Residual**: If outputs from the same source or different sources/databases can be combined to reveal information about a person or group of persons then there is a residual disclosure risk.

For each of our statistical and data releases, we will assess the risk of disclosure based on the following:

- Level of aggregation of the data;

- Number of tables produced from each dataset;

- Likelihood of an identification attempt;

- Size of the population;

- Consequences of disclosure;

To minimise the risk of disclosure and maximise the utility of the statistics in our statistical releases, we will use an appropriate combination of statistical disclosure control methodologies including: table design; rounding, primary suppression, and secondary suppression. We may use software for rounding and suppression purposes where feasible.

Details of the methods of rounding and suppression in use in the Department are included in Annex A.

Jude Hillary
Head of Profession for Statistics
Department for Education

ANNEX A

Practical Guide to the Application of Statistical Disclosure Control within the Statistical Policy on Confidentiality

Introduction

This document gives more detailed guidance on the disclosure controls applied within DfE statistical releases.

This policy replaces a range of different methods of disclosure control used in the department's statistical publications and related outputs up to 2010.

Much of DfE's published data is based on Census collections rather than surveys. This guidance relates primarily to the methods of rounding and suppression of figures from both sources for the purposes of maintaining confidentiality; separate decisions on rounding of the results of sample data to ensure representative reporting are out of scope of this guide.

Guidance

Primary suppression of small values (numerators), including zeros

When reporting on the number of children, pupils or teachers, a threshold of 3 is set for suppression so that values of 1 and 2 only are suppressed. This applies to all published statistics, with the exception of the following where a threshold of 6 is set so that values of 1 to 5 inclusive are suppressed:

- Looked After Children
- Private Fostering
- Children Accommodated in Secure Children's Homes
- Exclusions
- Sample survey data

Any further exemptions must be agreed with the Head of Profession.

When reporting on sensitive characteristics authors need to use professional judgement in when it might also be appropriate to suppress values relating to 0 or 100%. One example of this would be to suppress the entry if all the pupils in a school are eligible for free school meals.

The same judgement will be needed where rates close to 0 or 100% might reveal a small value for a sensitive characteristic. For example, if none of the pupils with a particular characteristic achieved 5 A*-C grade GCSEs, it might be possible to deduce from knowing one pupil's exam results that they do or do not have that characteristic, so suppression may be required.

Where a figure is suppressed, any rates calculated from that figure should also be suppressed.

Primary suppression of small values (denominators)

When calculating rates or percentages, those based on denominators of less than 3 will be suppressed. Exemptions to this are the same areas listed above in the numerators section. In these cases rates or percentages based on a population of 10 or fewer are suppressed. Survey data may require individual suppression thresholds to be set dependent on sample size.

In all cases, trailing decimals (i.e. when a cell within a spreadsheet is displayed to a set number of decimal places but contains a number to more decimal places) will be removed.

Secondary suppression and alternatives

Primary suppression alone is not sufficient disclosure control since suppressed figures can be deduced by looking at other figures in the table (e.g. sub-totals). There are three options open to DfE statistical release authors to overcome this:

- Secondary suppression (the suppression of the next lowest value in the row or column)
- Redesign of tables so the figures displayed are not small enough to be suppressed eg by combining columns or cells
- Rounding of subtotals and totals.

Statistical release authors will make every effort to use secondary suppression in the first instance, then redesign the tables, and then use rounding. Authors should be guided by transparency and choose the method which removes least utility from the statistics and leaves them closest to their original form.

Further rounding of figures, rates and percentages

Figures, rates and percentages may be rounded beyond the requirements for disclosure control where there is good reason to do so e.g. to avoid giving a false impression of accuracy with estimates, survey data, or where data quality issues exist; or to simplify very large numbers where there is little interest in the exact figure.

The decision to apply further rounding of figures will be made in the context of transparency and considering the views of users. Where a professional decision has been made that numbers need to be rounded beyond the requirements for disclosure control within a statistical release, those figures will be provided to a full level of detail within the underlying data supplied alongside the statistical release (i.e. only necessary disclosure controls

applied). The provision of underlying data has been required since July 2010 as part of the government's transparency agenda.

Use of symbols

The following approach is employed:

Symbol	Description
x	Suppressed for reasons of confidentiality
.. (double dots)	Not available eg where data was not returned for some reason
. (single dot)	Not applicable eg no schools of a particular type within an LA
- (dash)	Negligible count eg % below a certain level (usually 0.05%). Avoids showing very small numbers as zero.

Resources

From time to time, conflicts may arise between the needs of different groups of users, or between the needs of users and the resources available to meet those needs. In such circumstances it may be necessary to apply methods of disclosure control other than those specified in this annex. Such a decision should only be taken after assessment of the options and with referral to the Head of Profession. However, basic confidentiality protection must be maintained at all times.

Existing publications and previously published data

This policy became effective from December 2010. From this point, current years' data in statistical releases will follow these disclosure rules or, in complex cases, a plan to move to this basis will be in place. Each statistical release will clearly state what suppression and rounding has been applied and why, in both Technical Notes and the footnotes to each table. Back years in the time series within these publications do not need to be brought in line, though statistical authors may apply the rules retrospectively if they wish.