Department for Education

The National Pupil Database

2014-15 User Guide (UKDS Access)

Contents

Further information and contacts

Table of figures

Version control log

Version	Description	Author / date

1.   Introduction and background

1.1   Who is this guide for?

The National Pupil Database (NPD) forms a significant part of the evidence base for the education sector and supports accountability and school improvement.

Data from the NPD can be shared (under strict terms and conditions) with named bodies and third parties who, for the purpose of promoting the education or well-being of children in England, are:

- Conducting research or analysis;
- Producing statistics; or
- Providing information, advice or guidance.

This User Guide is aimed at those conducting research at UK academic institutions, (including that conducted as part of a taught degree) and at some research centres, and who are intending to apply for access to NPD data held at the UK Data Service at the University of Essex.

We advise all existing and potential users to read this User Guide before using the data or making a request to access it. **Note that the most sensitive data items can only be accessed by direct application to the Department for Education** (see https://www.gov.uk/national-pupil-database-apply-for-a-data-extract or contact npd.requests@education.gsi.gov.uk)**. The application process for these data is different from the UKDS process and users are advised to establish which items they require before embarking on completing an application.**

1.2   What is this guide for?

The department seeks to make information available to third parties subject to terms and conditions, while protecting the confidentiality of those individuals in the NPD dataset. It is important that the individuals who are described by the data are confident that their data is being:

- Collected and processed lawfully;
- Handled confidentially and with due care and respect for privacy; and
- Managed in accordance with relevant legislation e.g. Data Protection Act 1998, Section 1 of the Computer Misuse Act 1990 and Education (Information about Individual Pupils) (England) Regulations 2009.

The NPD User Guide:

- Describes what datasets and variables are held in the NPD.
- Describes the sensitivities of the NPD dataset.
- Provides guidelines for releasing NPD data.

1.3   What is the NPD?

The first version of the NPD, including information from the first pupil level School Census matched to attainment information, was produced in 2002.

The NPD is now one of the richest education datasets in the world holding a wide range of information about pupils and students and has provided invaluable evidence on educational performance to inform independent research, as well as studies commissioned by the department.

There are a range of data sources in the NPD providing information about children's education at different phases. The data includes detailed information about pupils' test and exam results, prior attainment and progression at each key stage for all state schools in England. The department also holds attainment data for pupils and students in non-maintained special schools, sixth form and further education colleges and (where available) independent schools.

The NPD also includes information about the characteristics of pupils in the state sector and non-maintained special schools such as their gender, ethnicity, first language, eligibility for free school meals, awarding of bursary funding for 16-19 year olds, information about special educational needs and detailed information about any absences and exclusions.

1.4   What data is in the NPD?

1.4.1 Sources of data in the NPD

The data held in the NPD is collected from a range of sources including schools, Local Authorities and awarding bodies. This data is processed by the department's Education Data Division and matched and stored in the NPD.

What data is held in the NPD for each academic year can be found in the NPD Data Tables Excel spreadsheet.

1.4.2 The School Census (Ages 2 - 19 plus)

Pupil level information was first collected in January 2002 as part of the Pupil Level Annual Schools Census (PLASC). The School Census replaced the PLASC in 2006 for secondary schools and in 2007 for nursery, primary and special schools.

The School Census is carried out three times a year in the spring, summer and autumn terms (January, May and October respectively) and provides the department with both pupil and school-level data.

Pupil level data is collected from the following types of schools:

- Nursery - maintained and direct grant nursery schools.
- Primary - middle deemed primary schools and academies.
- Secondary - middle deemed secondary schools and academies and City Technology Colleges.
- Special - maintained and non-maintained special schools, hospital special schools and academies.

Note: References to academies above include free schools.

The School Census is made up of separate modules. These modules collect information about:

- Pupil identifiers
- Pupil characteristics
- Pupil status
- Special educational needs
- Home addresses
- Absence
- Exclusions
- Post 16 learning aims (autumn only)

New and updated data from the spring, summer and autumn School Census collections is available in the NPD in July, August and January respectively.

1.4.3 PRU Census (Ages 2-19)

Pupil Referral Units (PRUs) are a type of school that are established and run by local authorities specifically for pupils who cannot attend a mainstream or special school. All PRUs have a Teacher in Charge (similar to a mainstream school's Head teacher) and a management committee which acts like a school's governing body. From 1st February 2008, all PRUs had to have a management committee that is established according to regulations and guidance.

The PRU Census is a mandatory data collection which was collected for the first time at pupil level in 2010. It is carried out once a year in January and collects data from all PRUs in England. The Census return can be completed by either the PRU itself or by the relevant local authority.

In January 2013, Alternative Provision (AP) academies and AP free schools completed the PRU Census.

New and updated data from the PRU Census is available in the NPD in July.

Note: From 2014 the PRU Census will be incorporated into the termly School Census.

1.4.4 Early Years Census (Ages 2, 3 and 4)

Since 2008, data has been collected for the Early Years Census at individual child level for 3 and 4 year olds (with a headcount total of 2 year olds). From 2014, individual child level data will also be collected for 2 year olds.

The Early Years Census is collected annually in January each year and collects data from all English providers of funded Early Years education including private, voluntary and independent (PVI) providers. The data collection is a snapshot of those children receiving funded Early Years education during the period of the week of the Census.

The Early Years Census only collects data about children taking up a free place. Any children who are not taking up a free place, or any Early Years provider that has no funded children at the time of the Census, are not included in the data collection.

New and updated data from the Early Years Census is available in the NPD in June.

1.4.5 Alternative Provision (Ages 2-19 plus)

The Census of pupils in Alternative Provision (AP) has been collected at an individual level since 2008. The Census is a local authority (LA) return that requires the submission of snapshot individual records of pupils in AP on Census day in January each year.

The Census covers pupils:

- Of school age attending a school not maintained by an authority for whom the LA is paying full tuition fees; and
- Educated otherwise than in schools and Pupil Referral Units under arrangements made by the LA.

There are a range of circumstances in which suitable education has to be provided for pupils outside school, including for those:

- Excluded from school on a permanent or fixed term basis;
- Unable to attend school for medical reasons; or
- Awaiting placement in a maintained school.

New and updated data from the Alternative Provision Census is available in the NPD in June.

1.4.6 Early Years Foundation Stage Profile (Age 4)

The Early Years Foundation Stage Profile (EYFSP) contains data about teacher assessments. Nurseries and schools submit the results of their EYFSP teacher assessments to local authorities at the end of the Summer Term. Local authorities provide the data to the department who then collate and validate the data before matching it into the NPD.

New and updated EYFSP data is available in the NPD in October.

1.4.7 Key Stage 1 (Age 5-7)

Key Stage 1 teacher assessments are undertaken by schools at the end of Year 2.

At Key Stage 1 the statutory subjects that all pupils must study are:

- Art and design
- Design and technology
- English
- Geography
- History
- Information and communication technology
- Mathematics
- Music
- Physical education
- Science

Religious education must also be provided.

Five main subjects are assessed at the end of Key Stage 1:

- Speaking and listening
- Reading
- Writing
- Mathematics
- Science

At Key Stage 1 the majority of pupils are expected to work at levels 1-3 in Key Stage and attain level 2 at the end of the Key Stage.

Schools submit the results of their Key Stage 1 teacher assessments to local authorities who provide the data to the department. The data is then collated and validated before being matched into the NPD.

The first version of new and updated Key Stage 1 teacher assessments is available in the NPD in September. The final version is available in November.

1.4.8 Key Stage 2 (Age 7-11)

Key Stage 2 tests and teacher assessments are taken at the end of Year 6.

At Key Stage 2, the statutory subjects all pupils must study are:

- Art and design
- Design and technology
- English
- Geography
- History
- Information and communication technology
- Mathematics
- Music
- Physical education
- Science

Religious education must also be provided.

The main subjects tested and assessed at the end of Key Stage 2 are:

- English
- Mathematics
- Science

At Key Stage 2 the majority of pupils are expected to work at levels 2-5 and attain level 4 at the end of the Key Stage.

Key Stage 2 tests

English and mathematics are currently tested at the end of Key Stage 2 (Science tests were discontinued in 2010). The English level is aggregated from separate reading and writing tests.

Teacher-assessed levels are also sent to the department and reported in Statistical First Releases (SFRs) and are reported at school level in the Performance Tables.

The data is supplied to the department by the Standards and Testing Agency. A contractor then creates datasets which are used in the creation of the Performance Tables and these are then matched into the NPD.

The first unamended Key Stage 2 data is available in the NPD in September.

The Performance Tables datasets are also sent to schools for checking. After schools have made any changes, the amended data is then matched into the NPD. The amended data is available in the NPD in December.

After the Performance Tables have been published, schools are given the opportunity to make errata changes and this new data is matched into the NPD. The final Key Stage 2 data becomes available in the NPD in March.

Key Stage 2 teacher assessments

Teacher assessments are conducted in English, mathematics and science.

Schools send the results of their Key Stage 2 teacher assessments to local authorities who provide the data to DfE. The department then collates and validates the data before matching it into the NPD.

1.4.9   Key Stage 3 (Age 11-14)

Key Stage 3 assessments are taken at the end of Year 9 and contain attainment data based on teacher assessments. Key Stage 3 tests were taken up to 2009.

At Key Stage 3, the statutory subjects that all pupils must study are:

- Art and design
- Citizenship
- Design and technology
- English
- Geography
- History
- Information and communication technology
- Mathematics
- Modern foreign languages
- Music
- Physical education
- Science

The teaching of sex education and religious education is also statutory.
Three main subjects are teacher-assessed at the end of Key Stage 3:

- English
- Mathematics
- Science

At Key Stage 3 the majority of pupils are expected to work at levels 3-7 and to attain level 5 or above at the end of the Key Stage.

Schools send the results of their Key Stage 3 teacher assessments to local authorities who provide the data to the department. The department collates and validates the data before matching it into the NPD.

Note: that the Department discontinued the collection of Key Stage 3 assessment data from 2013/14.

1.4.10 Key Stage 4 (Age 14-16)

Key Stage 4 examinations and assessments are usually taken at the end of Year 11. The NPD holds data about attainment in GCSE and equivalent qualifications taken at the end of Key Stage 4.

The Key Stage 4 data held in the NPD at individual subject level includes:

- Individual GCSE full courses
- GCSE short courses
- BTECs
- Applied Single and Double Award GCSEs

From 2010 the regulated Cambridge International Level 1 / 2 Certificates and their legacy iGCSE qualifications were included in the KS4 Performance Tables. Edexcel Level 1 / Level 2 certificates and their legacy iGCSEs were included from 2011. These qualifications are counted in the GCSE full and double award courses above.

Approximately 150 awarding bodies provide the department with separate data feeds of results data which are then matched from examination level to the individual student. When this has taken place, some additional indicators and flags are calculated and added to the data before matching into the NPD. The first 'unamended' Key Stage 4 data is available in the NPD in October.

The Performance Tables datasets are also sent to schools for checking. After schools have made any changes, the amended data is then matched into the NPD. The amended data is available in the NPD in January.

After the Performance Tables have been published, schools are given the opportunity to make errata changes and this new data is matched into the NPD. The final Key Stage 4 data becomes available in the NPD in April.

1.4.11 Key Stage 5 (Ages 16-18)

Key Stage 5 data contains data about attainment in A levels and equivalent qualifications.

The Key Stage 5 data held in the NPD at individual subject level includes:

- GCE A Level (including Applied GCE Single Awards).
- Applied GCE Double Awards.
- GCE AS Level (including Applied AS Level).
- Applied GCE A Level and AS combined.

The data includes students in maintained schools (including comprehensive schools, City Technology Colleges and academies, plus special schools where they have opted to provide results), independent schools and further education sector colleges.

Approximately 150 awarding bodies provide the department with separate data feeds of results data which are then matched from examination level to the individual student. When this has taken place, some additional indicators and flags are calculated and added to the data before matching into the NPD. After matching is completed, the first unamended Key Stage 5 data is available in the NPD in October.

The Performance Tables datasets are also sent to schools and colleges for checking. After schools have made any changes, the amended data is then matched into the NPD. The amended data is available in the NPD in January.

After the Performance Tables have been published, schools and colleges are given the opportunity to make errata changes and this new data is matched into the NPD. The final Key Stage 5 data becomes available in the NPD in April.

1.4.12 Children Looked After (Ages 0-18)

Information in the department's Children Looked After (CLA) database is collected at individual level and since 2005-06 has included the Unique Pupil Number. CLA data is collected annually between April and June for the previous financial year in the SSDA903 return.

This data is matched into the National Pupil Database so it is possible to analyse the attainment of children who are looked after.

New and updated CLA data is available in the NPD in September.

1.4.13 Children in Need (Ages 0-18)

The Children in Need Census captures child level information on children referred to and assessed by children's social care services within a set 12 month period (1st April - 31st March). Local authorities send data to the department between 1st April and 31st July.

The Census includes any child referred to children's social care services within the year and also any cases open at 1st April of the set year, for whom local authorities were committed to take an initiative. 'Taking an initiative' means:

- Active case work.
- Making regular payments.
- Where funding or on-going services such as respite care has been agreed.
- Maintaining a child with care and accommodation.
- A commitment to review the case at a predetermined date.
- Maintaining a child's name on a register that ensures the child and family receives targeted information or other special consideration.

New and updated Children in Need data are available in the NPD in October.

1.4.14 Post-16 Learning Aims - PLAMS (Ages 16-18)

Post-16 Learning Aims data is collected via the autumn School Census from secondary schools with sixth forms (including academies and City Technology Colleges). Data is not required from special schools, PRUs or secondary schools without sixth forms. The data must be submitted for pupils who are in National Curriculum year group 12 or above (regardless of their age and the level of their learning aims).

Note: Since 2006/07, the method of collecting learning aims data has changed. The quality of the data has improved and the resulting NPD extracts produced have altered in nature. Therefore, care should be taken when comparing learning aims across years, as the data may not be equivalent or consistent. Prior to 2012 learning aims were returned for Year 11 pupils. From 2012, learning aims have only been returned for those pupils in National Curriculum Year Group 12 or above.

Each pupil may have a number of learning aims, and for each one, six data items are collected:

- Qualification Accreditation Number
- Discount Code
- Learning Aim Start Date
- Learning Aim Planned End Date
- Learning Aim Actual End Date
- Learning Aim Completion Status

New and updated Post-16 Learning Aims data is available in the NPD at the end of January.

1.5. How is the data in the NPD used?

The NPD forms a significant part of the education evidence base and supports a number of key priorities around accountability and improvement for schools and children's services.

- The database supports schools' operational decisions by enabling them and inspectors to interrogate test and examination results in order to identify strengths and weaknesses and focus on those areas that need most improvement such as lesson planning and support for individual pupils.
- It supports local authority children's services operational decisions by enabling them to benchmark themselves against other authorities, in particular through the publication of the children in care and adoption Performance Tables and the Adoption scorecards.
- The database is the source for a wide range of analysis and statistics published in Statistical First Releases **here**. For example, each year the department publishes information about National Curriculum assessment and GCSE attainment by key pupil characteristics such as gender, ethnic group, special educational needs status and free school meals eligibility in England.
- The department uses the data to provide accurate target of funding for local authorities and schools, including the Pupil Premium and Revenue Support Grant. The data and statistical analyses are used to inform, influence and improve education and children's services policies and to monitor the performance of the education and children's services as a whole.
- The data held in the NPD is also used in a wide range of research including being linked, with consent, to a wide range of external datasets.

1.6   Who can the department share data with?

The department will not disclose pupils' and/or children's personal information without consent unless the law allows it to do so, and it is in compliance with the Data Protection Act.

The pupil level data held in the NPD is deemed 'personal' under the terms of the Data Protection Act 1998. The department has to make it clear to children and their parents what information is held about them, why it is held, the uses made of it by DfE and its partners. The department publishes a statement on its **website** setting out how the department processes pupils' and children's data.

The law allows the department to share pupil data from the NPD with named bodies and third parties including schools, local authorities, other government departments and agencies under the Education (Individual Pupil Information) (Prescribed Persons) (England) Regulations 2009, as amended.

These regulations also allow the department to disclose individual pupil information, subject to the Data Protection Act 1998, to named bodies and persons who, for the purpose of promoting the education or well-being of children in England are

- Conducting research or analysis;
- Producing statistics; or
- Providing information, advice or guidance.

Access is subject to requestors complying with terms and conditions imposed under contractual arrangements and a rigorous request approval process.

Selected data from the NPD may also be shared with the Office for National Statistics for the development of population and migration statistics and electoral registration officers to assist them in the development of electoral registers.

The department may decide to share pupil and children's information with third parties on a case by case basis where it is satisfied that to do so would be in accordance with the law and the Data Protection Act, and where it considers that such disclosure would promote the education or well-being of children.

1.7 The School Level Database (SLD)

1.7.1 What is the School Level Database?

This is a central database containing school level data, similar to that found in the NPD, and contains statistical information about schools in a single place.

The database contains a history of changes to schools since 2006 e.g. when schools opened and closed, changes in school type from maintained to academy and the details of predecessor and successor schools resulting from such changes.

The database contains reference data for a number of other types of institution as well as schools relevant to the data held in the database (e.g. Early Years Settings in the Early Years Foundation Stage Profile data and further education colleges in the Key Stage 5 data).

At the centre of the SLD is an institution history table (a spine) made up of schools by academic year from 2005/06 .The institution history table holds the following key identifiers:

- Unique Reference Number
- LA Number and Establishment Number
- Establishment Name

These are then linked to other contextual information such as:

- Establishment Postcode
- Establishment Type
- Gender of Establishment
- Phase of Education
- Whether the school is open or closed
- Links to predecessor and successor schools

The institution history table allows users of the data to link school level information from different sources together (e.g. School Census and Key Stage 4 Attainment), both within a year and across years.

Being able to track schools across years is important because their characteristics can change over time. For example,

- A change of a school's type as a result of the introduction of new policy e.g. academies.
- A new religious affiliation for a school.
- Changes in school organisation.

These changes may lead to the closure of some schools and the creation of new ones.

1.7.2 Where does the data come from?

Statistical information about schools comes from a range of different sources. Some is collected directly by the department e.g. school workforce data. Other data is collected on the department's behalf by contractors from examination and awarding bodies e.g. Key Stage 4 and 5 data. Some data is already provided at school level from source and some requires aggregating from pupil level (to a set of predefined business rules).

1.7.3 What data is in the SLD?

The SLD includes data from the following sources from 2006 (unless otherwise indicated):

- School Census
- Early Years Foundation Stage Profile (from 2007 only)
- Key Stage 1
- Key Stage 2
- Key Stage 3

- Key Stage 4
- Key Stage 5
- Level 3 Value Added
- Consistent Financial Reporting Schools Finance
- Schools Capacity (2009-2010)
- Geographical Data sourced from the ONS National Postcode Directory
- Institution History
- Institution Link

The School Level Database is updated as new data becomes available.

2. Requesting access to NPD data

2.1 What level of data is available?

Data from the NPD is made available at four different levels of detail. Table 1 shows the different level of data available and the tiers in which it is classified.

Table 1 - Information Tiers

Information Tier	Information Type	Description
1	Individual pupil level data - Identifying and / or Identifiable and Highly Sensitive **Note that this data is not available through the UKDS.**	Individual pupil level extracts that include identifying and highly sensitive information about pupils and their characteristics including items described as 'sensitive personal data' within the UK Data Protection Act 1998. Examples of identifying data items include names, address and date of birth. Examples of highly sensitive data items include looked after status, in need status, full ethnicity, full language and primary and secondary SEN Type, reasons for exclusions and absence and service child indicator.
2	Individual pupil level data - Identifiable and Sensitive	Individual pupil level extracts that include sensitive information about pupils and their characteristics including items described as 'sensitive personal data' within the UK Data Protection Act 1998 which have been recoded to become less sensitive. Examples of sensitive data items include ethnic group major, ethnic group minor, language group major, language group minor, special educational needs and eligibility for free school meals.
3	Aggregate School level data - Identifiable and Sensitive	Aggregated extracts of school level data from the department's School Level Database which include items described as 'sensitive personal data' within the Data Protection Act 1998 and could include small numbers and single counts. For example, there is one white boy eligible for free school meals in school x that did not achieve level 4 in English and maths at Key Stage 2.

4	Individual pupil level data - Identifiable	Individual pupil level extracts that do not contain information about pupils and their characteristics which is considered to be identifying or described as sensitive personal data within the Data Protection Act 1998. For example, the extracts may include information about pupil attainment, prior attainment, progression and pupil absences but do not include any trivially identifying data items like names and addresses and any information about pupil characteristics other than gender.

Tier 2 and 3 data can only be accessed in the safe room at the UKDS. Tier 4 data can be accessed over the virtual network. Anyone requesting access to sensitive data in tier 2 will be expected to explain on a field by field basis why each item is required and why the same outcome cannot be achieved using less sensitive data.

More information on the full list of data items available in each information tier is available in the **NPD data tables Excel** spreadsheet.

2.3 The approvals process

As the custodians of the National Pupil Database and other datasets held by DfE, a significant part of the role of the department's Education Data Division (EDD) is to process information. This includes the personal data of learners, children and of the school workforce for a range of purposes including research.

EDD's role is to ensure it continues to meet information governance standards in processing information and that this is lawful, secure and in line with departmental and governmental standards.

All applications for data made through UKDS will be assessed and approved / rejected by EDD at the Department for Education.

2.3.2 Purpose and role of the Data Request Team Approvals Process

The Data Request Team considers requests for access to data within tiers 2, 3 and 4 and undertakes the following activities:

a)   Establishes the credentials of the applicant(s) and the institution or organisation for which they work.

b)   Determines the names and job titles of all individuals who will be accessing the released data.

c)   Evaluates the intended uses for the data, and that these are consistent with those of the department and the reasons for which the data can be lawfully disclosed.

d)   Establishes with the applicant the minimum amount of information required to satisfy those purposes.

e)   Works with users to understand the nature of their requirements and establish whether there are conflicting needs.

f)   Establishes that the need for individual level data cannot be met by alternative means.

g)   Describes whether or not it discloses the micro-data requested (non-sensitive, sensitive, highly sensitive and identifying) and any conditions under which the information may become identifiable.

h)   Ensures that the requester has completed the data request application form

i)   Ensures that requests are technically sound, comply with the regulations which govern the release of pupil data, are practical and appropriate to the needs, and are ethical in nature.

j)   Informs applicants of the progress of their applications including decisions made by the Data Requests Team.

- Where applications are rejected, it will say why and what steps need to be taken for a resubmitted bid to be considered.

- Where applications are approved, it will keep applicants informed of progress towards meeting their request.

3. Sensitive Data

3.1 Pupil information

The data held within the NPD relates to the education and characteristics of individual pupils. Some information, such as that relating to ethnicity or special educational needs, could potentially be very sensitive in the sense that its disclosure could cause harm or distress to an individual pupil. While all personal information should be handled responsibly, some data items can be more sensitive than others or can more easily enable identification of individuals. For example:

The set of data items including names, address, postcode and date of birth (collectively referred to as 'the potential identifiers') could, in combination or with outside knowledge, help to identify an individual pupil;

Other reference numbers including the Unique Learner Number and Unique Pupil Number can uniquely identify a pupil if the relationship between the data item and the pupil is known;

Information relating to ethnicity and special educational needs is defined as sensitive personal data in the Data Protection Act 1998.

3.2 Are individuals identified?

The department has processes in place to make sure that sharing of identifying information is proportionate to need, and the identity of a pupil cannot be discovered using anonymised NPD data in isolation, even though potential identifiers (see 2.1) are held.

NPD Pupil IDs and Pupil Matching References are converted into anonymised Pupil Matching Reference (aPMR) numbers and added to the NPD dataset to reduce the dependence for access to the actual potential identifiers. This allows differentiation of pupils for comparative analysis, while still concealing the identities of individual pupils.

The aPMR is a data field in NPD. It was established to give each pupil in the NPD an identifier that is unique to them and the NPD. This was implemented to prevent much of the need for the disclosure of potential identifiers purely for matching purposes, although they may still be needed for matching with new data into the NPD. The aPMR can be used to link together records for a single pupil across years and data sources.

3.3 Tiers 1 and 2 data

The NPD includes data which is identifying and / or sensitive (including highly sensitive data). Table 2 provides details of identifying and sensitive data items in the NPD together with any corresponding mappings, where applicable. There will be circumstances where access to some or all of this data is required to undertake some research or analysis.

Table 2 - Identifying and Sensitive data items

Tier 1 data items (Note - these items are not available through UKDS)	Tier 2 data items
Pupil Matching Reference	Anonymised Pupil Matching Reference (aPMR)
Unique Pupil Numbers (UPN)	
Unique Learner Numbers (ULN)	
Exam candidate number	
Candidate ID	
Names	
Home address	Super Output Area (SOA) and other geographical areas
Postcode	
Output Area (OA)	
Date of birth / age within academic year	Age and month part of age of pupil at the start of the academic year. Year and month of birth
Ethnicity (extended ethnic codes)	Ethnic group (20 main group codes)
Language (extended language codes)	Language group (10 main group codes)
Primary and secondary special educational need (SEN) type	
Service child indicator	
Type of disability	
Pupil Premium indicator	
Children in Need (CIN) indicators	
Children Looked After (CLA) / In care indicators	
Reasons for absence	
Reasons for exclusion	
	SEN provision
	Free school meals (FSM) eligibility, including Ever FSM indicators
	Entry date
	Leaving date
	Boarder indicator
	Part-time indicator
	Pupil mobility indicator
	Distance to school indicators
	Persistent absentee indicator

Note: Tier 1 data items cannot be accessed through the UKDS. Requesters for this data should refer to https://www.gov.uk/national-pupil-database-apply-for-a-data-extract **or contact** npd.requests@education.gsi.gov.uk .

Anyone requesting access to sensitive data in tier 2 will be required to explain on a field by field basis why each item is required and why the same outcome cannot be achieved by using less sensitive data.

Note: The following data items (3.3.1 to 3.3.20) are considered to be identifying and highly sensitive and can only be accessed by direct application to the Department for Education (see https://www.gov.uk/national-pupil-database-apply-for-a-data-extract or contact npd.requests@education.gsi.gov.uk). **Requesters are expected to provide the rationale for requesting access to the data in their application.**

3.3.1   Pupil Matching Reference

The Pupil Matching Reference (PMR) uniquely identifies each pupil in the dataset. The PMR gives each pupil an identifier which is unique to them and allows matching across datasets without giving away their identity.

3.3.2   Unique Pupil Number (UPN)

UPNs were first introduced during the 1999/2000 academic year for learners in local authority maintained schools and, subsequently, academies in England (a similar compatible system has been introduced for learners in maintained schools in Wales). Independent schools are not required to issue UPNs for their learners, although many have done so on a voluntary basis.

UPN is a number that identifies each learner in England uniquely. UPNs are allocated to learners according to a nationally specified formula on first entry to school (or in some cases, earlier). Learners' UPNs are intended to remain with them throughout their school career regardless of any change in school or local authority (there are a few exceptional circumstances where this may not be the case).

3.3.3   Unique Learner Number (ULN)

The **Learning Records Service** assigns each young person a Unique Learner Number (ULN). This is a 10-digit reference number used alongside and to access the Personal Learning Record of anyone over the age of 14 involved in UK education or training.

The 10-digit ULN has been designed to ensure that no additional meaning can be inferred from its structure e.g. geographical location, level of learning. Its purpose is purely to provide a unique identifier for each individual's Personal Learning Record.

Learners will retain the same number for accessing their Personal Learning Record throughout their lives, whatever their level of learning and wherever they choose to participate in education, training and learning.

3.3.4 Exam candidate number

The Exam Candidate Number is the number given to the pupil by the relevant awarding body.

3.3.5 Candidate ID

The Candidate ID is the candidate serial number given by the relevant awarding body.

3.3.6 Pupil names

Surnames and forenames, which could include middle names, are collected in some of the datasets held in the NPD.

3.3.7 Pupil home address

Details of the home address of the pupil are collected in some datasets in the NPD.

3.3.8 Postcode

Geographic fields have been derived from postcode and are routinely available (see the worksheet named 'SLD-GEO NSPD-NSPL' in the **NPD data tables** Excel spreadsheet)

3.3.9 Census Output Areas

Output Areas (OAs) were devised by the Office for National Statistics to identify small geographical areas that cover similar population sizes, according to the 2011 Office for National Statistics Population Census. They are defined by adjacent postcodes as at Census day. There are 165,665 OAs in England and although too small for meaningful statistics in their own right, their main purpose is as building blocks for a range of other measures.

3.3.10 Pupil date of birth / age

The Date of Birth of a pupil can be aggregated up to:

- Age of pupil at the start of the academic year (in full years);
- Month part of age of the pupil at the start of the academic year;
- Year in which the pupil was born; or
- Month in which the pupil was born.

3.3.11 Ethnicity

Ethnicity is defined as 'sensitive personal data' under the UK Data Protection Act 1998.

Schools and Local Authorities are required to collect and record data on pupils' ethnic background - both for local purposes and to supply data to the department. The data enables the careers and experiences of children from different backgrounds to be monitored, and helps schools and Local Authorities to:

- Identify barriers to achievement.
- Establish strategies to raise standards.
- Comply with equal opportunities legislation and the Race Relations Amendment Act.
- Ensure effective allocation and targeting of funding.

The extended ethnicity codes can be re-coded by mapping to the 20 main ethnic groups and seven main ethnic categories and can be approved by the Education Data Division.

3.3.12 Language

The extended language codes can be re-coded by mapping to language groups and can be approved by the Education Data Division.

3.3.13 Primary and secondary Special Educational Needs

Primary Special Educational Need is the main type of need recorded for a pupil. Secondary Special Educational Needs is where a pupil has more than one Special Educational Need other than the main need recorded.

3.3.14 Service child indicator

The Service Children indicator recorded on the School Census indicates if a child has a parent or parents who are Service personnel in the regular HM Forces military units of all forces, or in the Armed Forces of another nation and stationed in England, and exercising parental care and responsibility.

3.3.15 Type of disability

The Equality Act states that someone is disabled if 'they have a mental or physical impairment which has a substantial and long-term adverse effect on their ability to carry out normal day-to-day activities'.

A Pupil's disability (including none and not collected) is grouped into particular types such as hearing, vision, mobility etc. The collection of disability was on a voluntary basis only and was only provided in the 2011 and 2012 January Census.

Note: This data item was removed from the School Census for 2013.

3.3.16 Pupil Premium

The Pupil Premium was introduced in April 2011 and is paid to local authorities by means of a specific grant based on pupil registration (from Reception to Year 11) as at the January School, Pupil Referral Unit and Alternative Provision Censuses each year.

The deprivation element of the Pupil Premium is allocated in respect of pupils who have been recorded as eligible for free school meals in any of the above pupil level Censuses in the last six years in both mainstream and non-mainstream settings.

The Pupil Premium is also payable in respect of those pupils recorded as children of Service personnel in any of the above pupil level Censuses from January 2011.

The Pupil Premium is paid to local authorities for those children who have been looked after continuously for at least six months, where at least some of the six months fell in the latest year ending 31st March. The premium is calculated using the SSDA903 returns which are collected each spring from all local authorities.

The Pupil Premium is also paid to academies via the Education Funding Agency.

3.3.17 Children Looked After (CLA)

The SSDA903 statistical return is an annual child-level data collection carried out by DfE. The return is completed for every child who was looked-after at any point during the year and also for a group of former looked-after children whose nineteenth birthday fell during the year.

Before 1998, the SSDA903 return covered all children looked after. Between 1998 and 2003, the return covered only a one third sample of children, those with a day of birth divisible by three. There was then an aggregate return (CLA 100) which was used to provide 'fast track' data and to gross the sample data. In 2000 the OC1 (educational qualifications of care leavers) return was introduced, followed by the AD1 (children adopted) returns in 2001 and the OC3 return (19th birthday) in 2002. From 2004, the

CLA100 return was discontinued as the SSDA903 reverted to covering all looked after children. Becoming a web based data collection for the first time, the SSDA903 return also incorporated the AD1, OC1 and OC3 returns.

3.3.18 Children in Need

The Children in Need census covers all children who are referred to Children's Social Care Services even if no further action is taken. This includes Children Looked After, those supported in their families or independently (CSF/I), and children who are the subject of a child protection plan.

Children in need are those assessed to be in need by children's social care services through an initial assessment.

3.3.19 Reasons for absence

Data on absences is collected in each Census and refers to the term prior to the Census date (i.e. the summer 2012 Census collects data on absences in the spring term).

Schools are able to provide absence data using a reason code or by using total figures for the number of half day sessions missed due to authorised or unauthorised absence. A small proportion of schools have provided both a breakdown of absence by reason and totals. In addition, a small number of schools do not have the required software to provide absence data by reason, and are therefore only able to provide overall totals.

Reason for absence was collected for the first time for the 2006/2007 academic year.

There are three attendance data releases each year as follows:

- May - autumn term
- October - autumn and spring terms
- March - full academic year

In 2013 the period for the collection of attendance data is being extended to the end of the summer term (31st July). In previous Census collections, attendance data for the summer term only covered up to the Sunday before the late Spring Bank Holiday.

From the start of the 2012/13 academic year the collection of attendance data will be extended to 4 year olds (but not for those attending nursery schools). Previously, attendance data was only collected for pupils aged 5 to 15.

3.3.20 Reasons for exclusion

Data on exclusions is collected in each Census and refers to the second term prior before the Census date (i.e. the summer 2012 Census collects data on exclusions in the Autumn 2011 term). There are two types of exclusions:

- A permanent exclusion refers to a pupil who is excluded and their name removed from the school roll. Such a pupil would then be educated at another school or via some other form of provision.

- A fixed period exclusion refers to a pupil who is excluded from a school but remains on the roll of that school because they are expected to return when the exclusion period is completed.

There is one exclusion data release each year (usually in July) which contains data on exclusions in the previous academic year.

For the 2005/06 school year, secondary schools also provided information on fixed period exclusions and the reasons for exclusion alongside the standard information they provided on permanent exclusions.

For the 2006/07 school year onwards, information on permanent and fixed period exclusions as well as the reasons for exclusion is available for primary, secondary and special schools.

Note: The following data items (3.3.21 to 3.3.29) are considered to be identifying and sensitive. Requesters are expected to provide the rationale for requesting access to this data in their application.

3.3.21 Special Educational Needs

Information about Special Educational Needs (SEN) is defined as 'sensitive personal data' under the Data Protection Act 1998.

The term 'special educational needs' has a legal definition as follows:

'Children with special educational needs all have learning difficulties or disabilities that make it harder for them to learn than most children of the same age. These children may need extra or different help from that given to other children of the same age.'

Special educational needs could mean that a child has difficulties with:

- All of the work in school.
- Reading, writing, number work or understanding information.
- Expressing themselves or understanding what others are saying.
- Making friends or relating to adults.
- Behaving properly in school.
- Organising themselves.
- Some kind of sensory or physical needs which may affect them in school.

3.3.22 Free School Meals

In England, children are eligible to receive Free School Meals (FSM) if their parents are in receipt of any of the following benefits:

- Income Support.
- Income-based Job Seekers' Allowance.
- Income-related Employment and Support Allowance.
- Support under Part VI of the Immigration and Asylum Act 1999.
- Guaranteed element of State Pension Credit.
- Child Tax Credit provided they are not also entitled to Working Tax Credit and have an annual gross income of no more than £16,190, as assessed by Her Majesty's Revenue and Customs.

Up to and including the 2012 Census collections, information was only returned on the current eligibility of pupils for FSM as at Census day.

From 2013 onwards, the collection of FSM eligibility information changed from a snapshot collection as at Census day to collection periods of eligibility between Censuses. For each period of eligibility a start date and end date will be collected.

3.2.23 Entry date

This represents the last date of entry of the pupil to the current school.

3.2.24 Leaving date

The date the pupil left the current school.

3.2.35 Boarder indicator

This indicates whether or not the pupil is a boarder at the school. The values that are returned in the Census for boarders are:

- Seven nights a week
- Six nights or less a week
- Nights per week not specified

3.2.26 Part-time indicator

This indicates whether a pupil is part-time in education (attends fewer than 10 sessions a week) or not. It does not indicate that a pupil is part-time at one or more establishments.

3.2.27 Pupil mobility indicator

This shows if the Pupil's postcode has changed since the previous Spring Census.

3.2.28 Distance to school indicators

These indicators show both the distance in miles from the pupil's postcode to their current school and to the nearest suitable school.

3.2.29 Persistent absentee indicator

The threshold at which a pupil is defined as '*persistently absent*' is 15 per cent of the number of possible sessions. This new threshold was introduced from the start of the 2011-12 academic year. In previous years the threshold was 20%.

4.   Guidelines on releasing NPD information

4.1   Fundamental principles of releasing NPD data

There are two basic principles to follow when releasing NPD data:

- Protecting the privacy and confidentiality of individuals; and
- Seeking advice from a member of the NPD Data Request team whenever there is any doubt about whether data can be released.

4.2   Published data

DfE publishes headline information from NPD data and makes the underlying data available to NPD users. Data can be requested but will not be made available to accredited users until key statistics using that data have been published.

The **statistical publication schedule** can be found on the DfE website.

4.3   Recognition of an individual and breaching confidentiality

NPD data users need to consider the risk of identifying individuals in their analyses prior to publication / release. You should be on the alert for any rare and justified and unintentional breach of confidence, such as responding to a query relating to a news item that may add more information to that already in the public domain.

Data should normally be released at a high enough level of aggregation to prevent others being able to 'recognise' a particular individual. If you recognise an individual while carrying out an analysis you must exercise professionalism and respect their confidentiality. If you believe this identification could easily be made by others you should alert a member of the NPD Data Request team within DfE for further advice by e-mail at **NPD.REQUESTS@education.gsi.gov.uk** Similarly, if you recognise an individual from existing analyses, you should alert the author of the analysis as well as the NPD Data Request team for further advice.

While appropriate handling of accidental recognition is acceptable, the consequences of deliberately breaching confidentiality could be severe, including criminal prosecution. Action may also be taken by the Department for Education against both your organisation and you as an individual.

4.4   Small numbers

Publishing small numbers of pupils displaying specific characteristics is not necessarily a problem when they cover a broad geographical area, because the pupil would not normally be identifiable. However, data that is likely to be more sensitive e.g. ethnicity should still be treated with care if it is likely to identify individuals. Small numbers within local authorities which include data considered as sensitive may allow identification of pupils and should not be published if the subsequent disclosure of the information could cause individuals harm and distress.

Statistical Disclosure Control

- Any published NPD figure(s) at a local level must be based on no fewer than 3 cases or else suppressed (e.g. replaced with an X, with a note "* in this table means the figures have been suppressed due to small numbers". Zero is allowed).
- Any published NPD figure(s) at a school level from 2011 onwards must be based on no fewer than 6 cases or else suppressed (e.g. replaced with an X and with a note "* in this table means a figure less than 6". Zero is allowed).
- Any published figures at school level for years prior to 2011 must be based on no fewer than 11 cases or else suppressed (e.g. replaced with an X and with a note "* in this table means a figure less than 11". Zero is allowed).
- Any analysis could enable the identification of individual pupils, especially with local knowledge. Even if only one cell requires cell suppression, you must suppress at least one other component cell (the next smallest) to avoid calculation of suppressed values from the totals.

The rules on suppression of low cell counts should be considered wherever small numbers are encountered, irrespective of whether the count is directly one of pupils.

Alternative methods can be considered to protect small numbers. For example, rounding or considering a higher level of aggregation for one or more items e.g. the move from postcode level to super output area (SOA), or from ethnic code level to ethnic group level. A higher level of aggregation might be the preferred option if several cells are affected by the suppression rule.

If these alternatives do not suit your needs, and you are unable to find a suitable compromise, contact the NPD Data Request team for advice at **NPD.REQUESTS@education.gsi.gov.uk.**

4.5   Footnotes

All analyses must carry the appropriate footnotes from this NPD User Guide concerning definitions, small numbers and data quality. It is particularly important that the data source (National Pupil Database, The Department for Education) is clearly referenced. These footnotes must be retained when any analyses are passed on, even if only part of the analysis is used.

5   Further information and contacts

We welcome your views on this User Guide and the data held in the NPD.

If you would like more information about the NPD, please contact our NPD Data Request team at **NPD.REQUESTS@education.gsi.gov.uk**.