



Department  
for Education

# **Alternative provision census 2017**

**Business and technical specification -  
version 1.1**

**January 2017**

# Contents

Version history	4
Assumptions made in creating this specification	6
1. Introduction	7
1.1 Purpose, scope and audience	7
1.2 Scope of alternative provision census	8
1.3 Structure of the alternative provision census	8
1.4 Changes from 2016 census	8
1.4.1 New data items	8
1.4.2 Existing data items	9
1.5 Pupil country of birth, pupil nationality and language – these data items are not required	9
2. Message header	10
3. Pupil/child level	11
3.1 Overall description and scope	11
3.2 Outline data content	12
3.2.1 Pupil / child identifiers module	13
3.2.2 Pupil/child characteristics module	13
3.2.3 Attendance module	17
3.2.4 Pupil / child SEN module	17
3.2.5 Home information module	17
3.2.6 Unique property reference number	18
3.3 Coverage and timing	18
4. Data return formats	19
4.1 Available formats	19
4.2 Data items with no values	19

4.3	Block entry	20
4.4	XML format submission	20
4.5	XML header	21
4.6	Pupil / child XML message structure	22
4.7	Special notes for XML returns	23
4.8	CSV format submission	24
4.9	CSV header	24
4.10	CSV pupil record structure (single pupil record)	26
4.11	CSV pupil record structure (two pupil records)	28
4.12	Special notes for CSV returns	33
5	Validation checks	34

## Version history

1.0	<p>All dates within the document updated to reflect 2017 collection period.</p> <p>1.4.1(a) – Language code – added</p> <p>1.4.1(b) – Pupil country of birth – added</p> <p>1.4.1(c) - Pupil nationality – added</p> <p>1.4.1(d) – Unique property reference number – added</p> <p>1.4.2(a) – Ethnicity – added</p> <p>1.4.2(b) – Adopted from care - added</p> <p>3.2.2(a) – Ethnicity – amended</p> <p>3.2.2(b) – Language code – added</p> <p>3.2.2(c) –Pupil country of birth – added</p> <p>3.2.2(d) – Pupil nationality – added</p> <p>3.2.2(f) – Universal infant free school meals – deleted – subsequent paragraphs renumbered</p> <p>3.2.2(g) – Post looked after arrangements – amended</p> <p>3.2.5 – Home information - amended</p> <p>3.2.6 – Unique property reference number – added</p> <p>4.2 – Data items with no value – amended</p> <p>4.6 – Pupil / child XML message structure – amended</p> <ul style="list-style-type: none"> <li>• Data item 100047 &lt;Language&gt; - added</li> <li>• Data item 100589 &lt;CountryofBirth&gt; - added</li> <li>• Data item 100588 &lt;Nationality&gt; - added</li> <li>• Data item 100587 &lt;PLAA&gt; - amended (formally &lt;AdoptedfromCare&gt;)</li> <li>• Data item 100586 &lt;UPRN&gt; - added</li> </ul> <p>4.10 – CSV pupil record structure (single pupil) –</p>	Phil Dent 15/04/2016
-----	---	-------------------------

	<p>amended</p> <p>4.11 – CSV pupil record structure (two pupils) - amended</p>	
1.1	<p>1.4.1(a) – Language code – deleted</p> <p>1.4.1(b) – Pupil country of birth – deleted</p> <p>1.4.1(c) - Pupil nationality – deleted</p> <p>1.5 – Pupil country of birth, pupil nationality and language – section added to explain data not required</p> <p>3.2.2(b) – Language code – deleted</p> <p>3.2.2(c) –Pupil country of birth – deleted</p> <p>3.2.2(d) – Pupil nationality – deleted</p> <p>4.6 – Pupil / child XML message structure – amended</p> <ul style="list-style-type: none"> <li>• Data item 100047 &lt;Language&gt; - deleted</li> <li>• Data item 100589 &lt;CountryofBirth&gt; - deleted</li> <li>• Data item 100588 &lt;Nationality&gt; - deleted</li> </ul> <p>4.10 – CSV pupil record structure (single pupil) – amended – to delete above data items</p> <p>4.11 – CSV pupil record structure (two pupils) – amended – to delete above data items</p> <p>Data item 100586 &lt;UPRN&gt; amended to &lt;UniquePropertyReferenceNumber&gt; to comply with CBDS</p>	<p>Phil Dent 22/11/2016</p>

This specification must be read in conjunction with the Excel workbook giving the current version of the revised [Common Basic Data Set \(CBDS\) database](#)

# Assumptions made in creating this specification

A1	The census will use COLLECT as the data collection mechanism.
A2	Education Data Division will only accept submissions in XML format. A CSV to XML converter will be provided on request to convert fixed format CSV data into the required XML format.

# 1. Introduction

## 1.1 Purpose, scope and audience

This specification describes the 2017 census of pupils in alternative provision (AP) which will enable local authorities, and those such as software suppliers working on their behalf, to prepare the necessary data and processes for compliance.

The specification covers the expected census return to department and the validation rules which will be applied to the data submitted.

The following diagram outlines the overall scope of the collection:

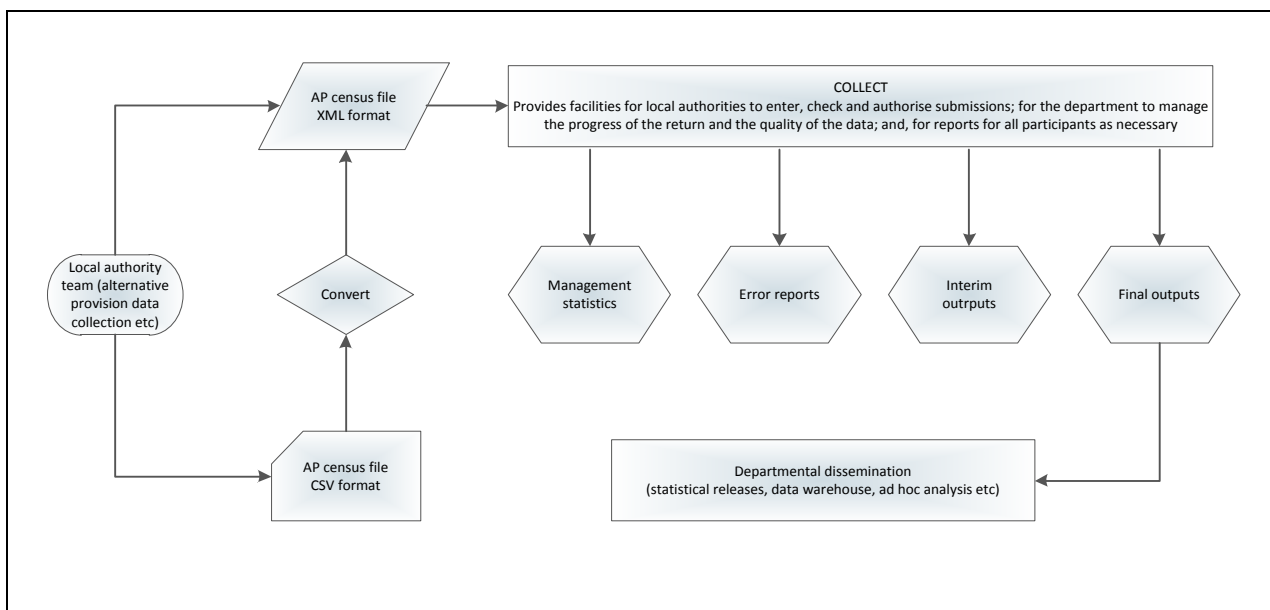


Figure 1 - High level data collection process

This requirements specification consists of:

- This Word document containing a narrative, sample XML and CSV messages
- The validations Excel workbook
- The [CBDS](#) Excel workbook
- The XML schemas will be supplied separately and will also be published on the department's [website](#) when available.

## 1.2 Scope of alternative provision census

The census of pupils in alternative provision is a local authority census that requires the return of individual pupil records. The census covers pupils attending a school not maintained by a local authority for whom the authority is paying full tuition fees, or educated otherwise than in schools and pupil referral units, under arrangements made (and funded) by the authority.

Pupils should not be included in the AP census if they are included in one of the following census:

- the school census for maintained nursery, primary, secondary, all-through schools (including academies and free schools) and maintained, non-maintained and hospital special schools, and PRUs (including AP academies and AP free schools) in England
- the annual school census for maintained general hospital schools
- the early years census

Pupils at non-maintained special schools may fall within the scope of the AP census where they are local authority funded (see [paragraph 3.1\(c\)](#) below for details).

The annual AP census has a 'census date' of the third Thursday in January. Unless otherwise stated, the terms 'census date' and 'census day' in this document refer to 19 January 2017.

## 1.3 Structure of the alternative provision census

The 2017 AP census consists of pupil / child level data. This is comprised of one or more modules (each module being a subset of the pupil / child level data).

## 1.4 Changes from 2016 census

### 1.4.1 New data items

#### (a) Unique property reference number

This new data item is added to the home information module and allows the unique property reference number (100586) to be returned where available, see [paragraph 3.2.6](#) for full details.



## **1.4.2 Existing data items**

### **(a) Ethnicity**

The age restriction which previously only made the collection of ethnicity (100563) mandatory for pupils aged 5 or over has now been removed. As such, this data item now applies to all pupils in scope of the AP census.

### **(b) Adopted from care**

The 'adopted from care' name for this data item does not provide an accurate description of the data which is expected to be recorded within it. This data item is intended to record 4 variants of children ceasing to be looked after (3 of which do not relate to adoption) and therefore avoid confusion the name of this data item has been changed from 'Adopted from Care' (100549) to 'Post looked after arrangements' (100587).

## **1.5 Pupil country of birth, pupil nationality and language – these data items are not required**

We are not extending the collection of pupil country of birth and pupil nationality data to the alternative provision census nor will be collecting information on language.

For suppliers who are not able to remove these fields from their extraction routine for the 2017 collection, the department will perform a file transformation during the loading process into COLLECT to remove the fields not required from the return and ensure that this data is not returned to the department.

## 2. Message header

Section 4 describes two formats for the return of the data - XML or CSV. For either format, the data items required within the header are as follows:

<p>Survey collection name &lt;Collection&gt; (800001) will be 'APC'          Description: Name of the data collection eg APC</p>
<p>Survey term &lt;Term&gt; (800002) will be 'SPR'          Description: SPR denotes that the collection is carried out in the spring term</p>
<p>Survey year &lt;Year&gt; (800003) will be '2017'          Description: The year of the collection</p>
<p>Survey reference date &lt;ReferenceDate&gt; (800004) will be '2017-01-19'          Description: The reference date is normally the day of the census or collection and is used as the baseline date from which comparisons with other dates in the return can be made</p>
<p>Source level &lt;SourceLevel&gt; (800005) will be 'L'          Description: Set to L for a local authority system</p>
<p>LA &lt;LEA&gt; (200001) is the LA number          Description: Standard DfE three digit local authority number</p>
<p>Software code &lt;SoftwareCode&gt; (800006) a free format field that can be populated with whatever identifier the software supplier wishes to use          Description: Software supplier reference</p>
<p>Release &lt;Release&gt;,(800008) a free format field that can be populated with whatever identifier the software supplier wishes to use to identify the release          Description: A code/date (or combination) provided by the software supplier to assist in identification of the release</p>
<p>Serial No &lt;SerialNo&gt; (800007) will start at '001' and then be incremented by 1 each time a new file is prepared for submission          Description: An incremented number, generated by a provider's software, enables data collection systems to identify re-submissions</p>
<p>Date / time &lt;DateTime&gt; (800010) is the date and time when the file was prepared or generated in CCYY-MM-DDThh:mm:ss format          Description: Date and time of generation of the return</p>
<p>CBDS level &lt;CBDSLevel&gt; (800011) 'Pupil' should be provided          Description: Shows which of the defined CBDS levels are present within a data return</p>

## 3. Pupil/child level

### 3.1 Overall description and scope

This section describes the pupil/child level element of the 2017 AP census.

Pupils in scope are classified according to the type of alternative provision, which should be one of the following:

- a) Independent school
- b) Hospital
- c) Non-maintained special school
- d) Not a school

**(a) 'Independent school'** refers to independent schools in England and Wales for whom the local authority is paying full or part tuition fees. This includes pupils attending independent schools within an authority and in another area for whom the full cost of tuition is paid for by the local authority wholly or in conjunction with social services and / or the health authorities. Do **NOT** include 'imported pupils' ie those pupils attending an independent school in a local authority's area for whom another local authority is paying tuition fees.

**(b) 'Hospital'** refers to children who are in hospital and where the hospital is not a hospital school. For "hospital" and "non-maintained special schools" please include "exported" pupils, ie those attending an establishment in another authority's area for whom your local authority is paying tuition fees. "Imported" pupils (ie those attending an establishment in your area for whom another local authority is paying the fees) are **excluded** from the return.

**(c) 'Non-maintained special schools'** comprises those pupils attending such schools in your area or another authority's area for whom your authority is paying full or part tuition fees. This includes those pupils for whom the full cost of tuition is paid for by the authority wholly or in conjunction with social services and / or the health authorities. Pupils attending an establishment in your area for whom another authority is paying the fees are **excluded** from the return.

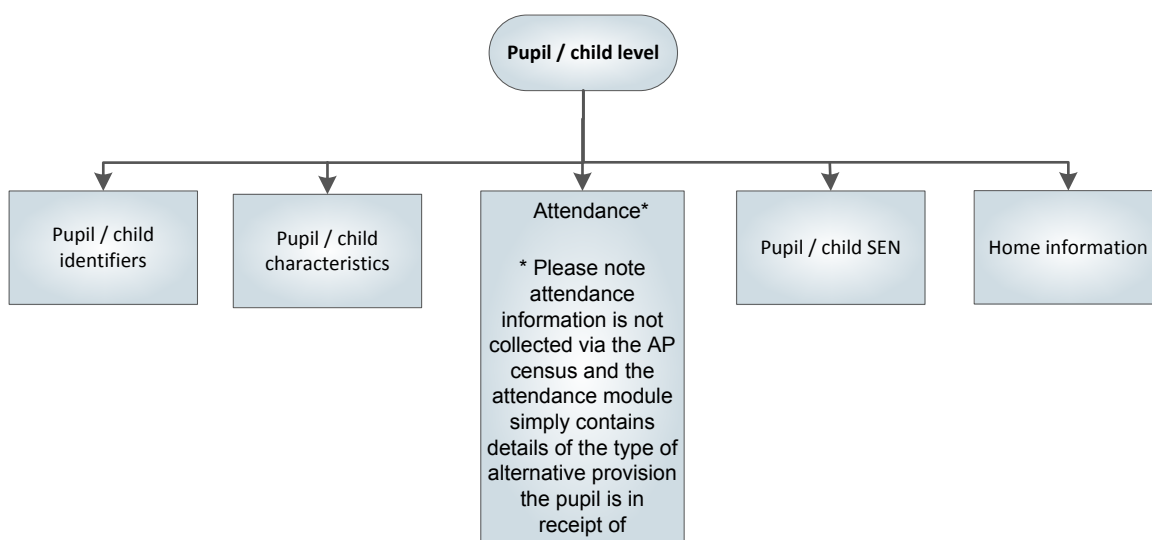
**(d) 'Not a school'** covers pupils being educated otherwise than at school. This includes the following groups:

- those educated in community homes or units (include pupils in independent residential homes - located either inside or outside your local authority area - where education is provided on site and for whom the authority is wholly responsible for their educational provision) and where the pupil is not registered as attending a maintained school

- pupils taught at home - only include those pupils who are receiving local authority funding. Children whose parents have decided to educate them at home are **not** included in the return, unless the local authority is providing significant support to enable the child to attend a college of further education or another alternative provider, or to support their special educational needs
- arrangements made for the education of pupils in travellers' families. This category does **not** include those traveller pupils who are on roll at a school on census day
- pupils of compulsory school age (not registered at a school) who are being educated at further education colleges and for whom the local authority are financially responsible
- asylum seekers of compulsory school age attending further education colleges or with voluntary sector providers

## 3.2 Outline data content

The return contains five modules for each pupil as in Figure 2 below:



**Figure 2 – Pupil / child level modules**

Full descriptions of the data items in each of these modules are shown in the common basic data set ([CBDS](#))

The paragraphs below contain guidance.

Section 4.6 contains a sample XML message structure which shows the modular layout.

### 3.2.1 Pupil / child identifiers module

This module contains the pupil's unique pupil number (UPN), name, date of birth and gender.

If the pupil has a UPN (many pupils in alternative provision have previously been in a mainstream school) then it should be provided. If a pupil has not previously been assigned a UPN then local authorities are asked to generate one.

#### Unique learner number

Unique learner numbers (ULNs) are assigned to all students over the age of 14 in publicly funded education and training and the ULN are mandatory for use in English schools. They are allocated and managed by the [Learning Records Service](#) (LRS) and will remain with the individual for their lifetime. If a student does not already have a ULN assigned to them, local authorities are required to sign the learning provider agreement which will provide local authorities with access the LRS system which will enable them to issue and manage ULNs **only** for pupils receiving alternative provision and who are not associated with a school having a UKPRN and access of its own to system. These users will need to complete an [Organisation Registration form](#) to nominate a super user for the system.

### 3.2.2 Pupil/child characteristics module

#### (a) Ethnicity

This module contains the pupil's ethnicity code (100563). The [CBDS](#) code set (D00250) provided for the return contains the department's main codes and extended codes - these are the same as those specified for the school census. Local authorities that use extended codes for pupils in schools should also use them for alternative provision. Conversely, those that use only the main codes in schools should also use only the main codes for alternative provision.

The age restriction which previously only made the collection of ethnicity mandatory for pupils aged 5 or over has now been removed. As such, this data item now applies to all pupils in scope of the AP census.

#### (b) Free school meal eligibility

The provision of additional funding via the pupil premium directs extra funding to those children from deprived backgrounds and is based on their eligibility for free school meals (FSM) at any time in the last six years - known as 'FSM ever'.

Pupils may only be recorded as FSM eligible if they meet the FSM eligibility criteria (ie in respect to family income) and make a claim. Pupils who are **only** in receipt of a free school lunch due to a universal entitlement - such as the infant pupil universal entitlement – are **not** recorded as FSM eligible and therefore are not eligible to receive pupil premium.

The actual allocation of the pupil premium for 2017-18 for pupils in alternative provision will be based on pupils aged 4 to 15 as at 31 August 2016 recorded in the 2017 alternative provision census with an AP type of independent, non-maintained special school and hospital and pupils aged 4 to 15 with an AP type of not in school who were recorded as known to be eligible for free school meals on any school census, PRU Census or AP census since summer 2011.

The department will determine eligibility for the pupil premium by matching the 2017 AP census to previous census collections in the national pupil database (NPD).

The pupil premium only covers periods of FSM eligibility experienced in England. Common transfer files (CTFs) received from schools or local authorities outside of England (eg Wales) when children transfer to AP in England may include FSM eligibility whilst the pupil was at a school or setting outside of England. This period of FSM eligibility in a school or AP outside of England does not attract funding via the pupil premium so it is important that the FSM eligibility end date reflects the date the pupil left the school or AP outside of England and that the appropriate country code eg Wales, is associated with the period of FSM eligibility. This will enable the department to exclude any pupils from pupil premium funding who have only experienced FSM eligibility in any country other than England and for local authorities to see in their systems that these pupils will not attract funding.

FSM eligibility start date (100484), FSM eligibility end date (100485) and Country of UK (200634) will be collected for pupils within the scope of the collection, who are on roll on census day, and have periods of FSM eligibility since the last census ie those with:

- a) An FSM eligibility start date and a blank FSM eligibility end date (ie eligibility ongoing on census day 19 January 2017); or
- b) An FSM eligibility end date since the previous census ie on or after 21 January 2016, and on or before the current census day, 19 January 2017.

The data returned reflects the position as at census day. Where a system holds a period of FSM eligibility that matches the criteria for the return but that period has a stored FSM eligibility end date that is after the census day, then the returned version should include an FSM eligibility start date but no FSM eligibility end date. Similarly FSM eligibility start dates returned in the census should be on or before census day, not after it. Stored start or end dates after census day could not have been known on census day and so should not be returned.

Multiple FSM eligibility start and end dates falling within the period should be returned. For example if a pupil had been eligible for FSM from 1 September 2014 until 31 October 2016 and then became eligible again from 1 December 2016 then the following would be returned:

For the first period of eligibility:

FSM eligibility start date of 2014-09-01 and FSM eligibility end date of 2016-10-31; and,

For the second period of eligibility:

FSM eligibility start date of 2016-12-01 and no FSM eligibility end date

Each period of eligibility should have an associated country of UK attached to enable the department to ensure that those pupils who have only experienced periods of FSM eligibility outside of England do not attract the pupil premium.

### **(c) Service children in education**

The 'service children in education indicator' data item indicates if a child has a parent or parents who are classified as category 1 or 2 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.

Responses will be 'Yes', 'No' and 'Refused' with an additional code of 'Unknown' to be used to indicate no response given or other reason for no information.

Note that such children would not normally be expected to be eligible for free school meals and this is reflected in the validation.

### **(d) Post looked after arrangements**

Post looked after arrangements (formally adopted from care) <PLAA> (100587) identifies those children on roll on census day who were looked after (in England or Wales) immediately before adoption or prior to being subject of a child arrangement order (previously known as a residence order) or special guardianship order. It will be up to those who have parental responsibility (adoptive parents, special guardians and carers) to decide if they wish the local authority / provider to know their child's status and provide the necessary evidence, for example, a copy of the adoption order. The local authority / provider will be able to record the information using the following codeset; only one of the values will apply to each pupil:

<b>Code</b>	<b>Post looked after arrangements</b>
N	Not declared
A	Ceased to be looked after through adoption
G	Ceased to be looked after through a special guardianship order (SGO)
R	Ceased to be looked after through a residence order (RO)
C	Ceased to be looked after through a child arrangement order(CAO)

The default value will be N, not declared.

A [template letter](#) is available for local authorities to adapt and send out to their providers to help assist with the identification and allocation of the pupil premium to pupils in

alternative provision who have left care under an adoption, special guardianship or residential / child arrangements order.

### **(e) Early years pupil premium**

The [early years pupil premium \(EYPP\)](#) is additional funding for early years settings to improve the education they provide for disadvantaged three and four year olds.

The data collected via the AP census will be used by the department to determine the early years pupil premium allocation which the department provides to local authorities to fund their early years providers.

Early years pupil premium data is required from all children aged 3 included on the AP census (ie those born between 01/09/2012 and 31/12/2013 (inclusive)). Children aged 4 at the start of the academic year are eligible for the school age pupil premium and therefore are not eligible for early years pupil premium funding.

Children will be eligible for EYPP, and should therefore be recorded with early years pupil premium eligibility equal to TRUE, if they are receiving any hours of government funded early education and either::

- meet the benefits related criteria for Free School Meals (please note: meals delivered as part of the universal entitlement are **not** FSM)
- are in the care of the local authority (in England or Wales)
- have left care (in England or Wales) through:
  - adoption
  - special guardianship
  - a child arrangement order (formally known as a residence order)

For those pupils eligible for the early years pupil premium, the reason why they are eligible is also required against the basis of eligibility field (100560). This should be recorded as follows:

- [EE] – eligible through economic reasons: where they are eligible via the benefits related criteria for FSM
- [EO] – eligible through other reasons: where they are eligible due to being in care or due to leaving care through adoption, a special guardianship order or a child arrangement order
- [EB] – eligible through both reasons: where they are eligible through both economic and other reasons
- [EU] – eligible through unknown basis: where the school or AP provider knows the child is eligible for EYPP (due to receiving funding from the local authority) but does not necessarily know the reason why they are eligible



### 3.2.3 Attendance module

This module contains the type of alternative provision that is being provided.

### 3.2.4 Pupil / child SEN module

The 2014 [code of practice](#) removed the requirement for separate 'school action or early years action' and 'school action plus or early years action plus' categories – these were replaced by a single 'SEN support' category (code 'K'). [The Children and Families Act 2014](#) replaced 'SEN statements' with 'Education, Health and Care plans' (EHC plans).

During the transitional period - to April 2018 - providers will be able to record pupils as having either 'S' (statement) or 'E' (EHC plan). Providers may have some pupils with statements and some who have already moved to EHC plans but no individual pupil can have both. An EHC plan is not just the rebadging of a statement. To ensure that EHC plans are person-centred and focused on outcomes, an EHC plan must only be issued following a 'transfer review' – an EHC needs assessment. For further information see the [statutory guidance](#) chapter 6.

Code 'S' may be retained for existing pupils up to April - no new statements can be issued.

### 3.2.5 Home information module

This module gives information about where a pupil lives **during term time**, and will be a **UK address** (In some cases this may be the school address).

Address data should only be provided for current addresses (ie where pupil address type (100102) = C). Multiple current addresses should be provided where available.

Post code (100121) must be provided together with the remainder of the address in either BS7666 Version 1.4 or line address format. The unique property reference number (100586) should also be provided where available.

Depending on the nature of the address not all items are necessary - for example an address in line address format might only consist of two out of the five line address items. For addresses in BS7666 format the PAON and street must be present and at least one of: locality; town; administrative area; or, post town.

BS7666 Format items:

- SAON (secondary addressable object name) (100103)
- PAON (primary addressable object name) (100109)
- street (100115)
- locality (100116)
- town (100117)

- administrative area (100118)
- post town (100119)

Or, line address format Items:

- address line 1 (100128)
- address line 2 (100129)
- address line 3 (100130)
- address line 4 (100131)
- address line 5 (100132)

The department recognises that schools may not have enough time to convert all their pupils' addresses to BS7666 Version 1.4 format. Address details can be supplied in either BS7666 Version 1.4 format or line address format, with postcode (and unique property reference number where available) to be supplied regardless of format used.

### **3.2.6 Unique property reference number**

The unique property reference number (100586) ['UPRN']; introduced on a voluntary basis from 2017; is defined by Ordnance Survey as: 'The unique identifier for every spatial address in Great Britain. It provides a comprehensive, complete, consistent identifier throughout a property's life cycle – from planning permission through to demolition'.

Where available within local authority systems, the 'UPRN' should be returned alongside the post code and full address in either BS7666 or address line format. However where UPRN is not populated within systems the data is not required.

### **3.3 Coverage and timing**

Pupils should be aged between 3 (as at 31 December) and 18 (at 31 August) ie those pupils born between 01/09/1997 and 31/12/2013. However, pupils aged 2 (as at 31 December 2016) can be included if they are funded by the local authority.

## 4. Data return formats

**Please note: example data shown in this section is for illustration purposes only.**

### 4.1 Available formats

The alternative provision census is submitted to the department in XML format, as described in this section.

For local authorities without a MIS that can output an AP census compliant XML file, the department provides a data entry spreadsheet which enables the input of file header and individual pupil record information. On completion the file is converted to XML format (using the spreadsheet) and uploaded to COLLECT. The data entry spreadsheet is provided on request.

### 4.2 Data items with no values

There are three scenarios when there may be no values for particular data items. These are:

- Standardised header information that is not relevant to this census. The only data item that should be not be present is:
  - establishment
- Values contingent on the content of other data items. These items are:
  - home address details (see paragraph [3.2.5](#))
  - free school meal (FSM) eligibility information is also not required for pupils with AP type of 'NOT' (not in school)
- Values which are described in this specification as optional, for example. These items are:
  - pupil / child middle names
  - pupil / child preferred surname
  - unique property reference number

In the above scenarios the XML tag is not required to be present in the census return. For all other scenarios, a missing item or value will result in an error being generated.

Within XML, an optional item that is omitted would not show a tag. Within CSV, an optional item that is omitted is shown by a new comma separator.

Instances may arise where all of the data items within an XML container have no value. In such cases, the XML container is not required to be present in the return. For

example, if a pupil has no periods of FSM eligibility then as well as the <FSMstartDate>, <FSMendDate> and <UKcountry> not being required, <FSMperiod> and <FSMeligibility> would also not be present.

### 4.3 Block entry

For suppliers creating software applications a block entry facility for fields is highly desirable in order to save time. This facility allows the user to choose a selection of pupils (eg a type of provision) and specify a value to be applied to all. However, it must also allow data to be edited on an individual basis.

### 4.4 XML format submission

An XML submission is a single message consisting of:

- A header
- A repeating group containing the data for each pupil

Overall message structure
<Message>
<Header> - see 4.5 below for details
<APpupilsChildren> - see 4.6 below for details
</Message>

## 4.5 XML header

A standard XML message header is required for each census file. The data items should be populated according to the guidance in Section 2. The layout is as follows:

Header structure	
	<Header>
	<CollectionDetails>
800001	<Collection>APC</Collection>
800002	<Term>SPR</Term>
800003	<Year>2017</Year>
800004	<ReferenceDate>2017-01-19</ReferenceDate>
	</CollectionDetails>
	<Source>
800005	<SourceLevel>L</SourceLevel>
200001	<LEA>999</LEA>
800006	<SoftwareCode>Software Supplier Reference</SoftwareCode>
800008	<Release>1.4 2016-10-11</Release>
800007	<SerialNo>001</SerialNo>
800010	<DateTime>2017-01-19T11:14:05</DateTime>
	</Source>
	<Content>
	<CBDSLevels>
800011	<CBDSLevel>Pupil</CBDSLevel>
	</CBDSLevels>
	</Content>
	</Header>

(Please note that the <Source><Estab> item, used in some DfE collections, is not used here and is omitted).

## 4.6 Pupil / child XML message structure

Pupil / child XML message structure		Cardinality (Note: item level elements occur only once unless otherwise stated)
	<APpupilsChildren>	
	<PupilChild>	
	<PupilChildIdentifiers>	
100001	<UPN>X1234567898765</UPN>	
100016	<UniqueLearnerNumber>1234567890</UniqueLearnerNumber>	
100003	<Surname>Familyname</Surname>	
100004	<Forename>Firstname</Forename>	
100006	<MiddleNames>Middlename</MiddleNames>	
100011	<PreferredSurname>Familyname</PreferredSurname>	
100292	<PersonBirthDate>1066-03-24</PersonBirthDate>	
100287	<GenderCurrent>2</GenderCurrent>	
	</PupilChildIdentifiers>	
	<PupilChildCharacteristics>	
100563	<Ethnicity>WBRI</Ethnicity>	
	<FSMeligibility>	Each <PupilChildCharacteristics> group contains 0..1 <FSMeligibility> group
	<FSMperiod>	Each <FSMeligibility> group contains 1..n <FSMperiod> groups
100484	<FSMstartDate>1087-10-09</FSMstartDate>	
100485	<FSMendDate>1087-11-16</FSMendDate>	May be missing in one and only one <FSMperiod> group
200634	<UKcountry>ENG</UKcountry>	
	</FSMperiod>	
	</FSMeligibility>	
100559	<EYPPE>Y</EYPPE>	Applies all pupils aged 3 ONLY
100560	<EYPPBF>EE</EYPPBF>	Applies all pupils aged 3 ONLY
100330	<ServiceChild>N</ServiceChild>	
100587	<PLAA>N</PLAA>	
	</PupilChildCharacteristics>	
	<Attendance>	
100471	<APTtype>NOT</APTtype>	
	</Attendance>	

	<SpecialEducationalNeeds>	
100580	<SENprovision>N</SENprovision>	
	</SpecialEducationalNeeds>	
	<HomeInformation>	
100121	<PostCode>ZZ99 9XX</PostCode>	
100586	<UniquePropertyReferenceNumber>123456789102</UniquePropertyReferenceNumber>	
	<BS7666Format>	
100103	<SAON> 2 Flat</SAON>	
100109	<PAON> 123</PAON>	
100115	<Street>Street Lane</Street>	
100116	<Locality>Localarea</Locality>	
100117	<Town>Townbury</Town>	
100118	<AdministrativeArea></AdministrativeArea>	
100119	<PostTown>Postaltown</PostTown>	
	</BS7666Format>	
	<LineAddressFormat>	
100128	<AddressLine1>123 Street Lane</AddressLine1>	
100129	<AddressLine2>Localarea</AddressLine2>	
100130	<AddressLine3>Townbury</AddressLine3>	
100131	<AddressLine4>Postaltown</AddressLine4>	
100132	<AddressLine5>Countyshire</AddressLine5>	
	</LineAddressFormat>	
	</HomeInformation>	
	</PupilChild>	
	</APpupilsChildren>	

## 4.7 Special notes for XML returns

The standard XML entity references should be used for the following special characters:

Character	Entity reference
Ampersand (&)	&amp;
Left angle bracket (<)	&lt;
Right angle bracket (>)	&gt;
Single quote / apostrophe (')	&apos;
Double quotes (")	&quot;

For special characters such as é use a character reference such as '&#233'. This produces a generic XML file which can be viewed as such in a suitable browser. Details of all such characters are found within the XML standard documentation - see for example [Extensible Markup Language \(XML\) 1.0 \(Fifth Edition\)](#) Section 2.2.

The department recommends that the file is given a meaningful name that includes the local authority number (the 3 digit serial number). The file extension is '.xml'. This will help users in selecting the correct file to upload to COLLECT.

## 4.8 CSV format submission

As an alternative to XML, local authorities may produce a comma separate value format (CSV). Please note that the e-Government interoperability framework encourages the use of XML for data transfer. XML is the department's format of choice and is preferred for future proofing. For those local authorities who have not developed XML facilities, a CSV file may be produced and converted using the CSV to XML converter which is supplied on request.

A CSV format data file is an ASCII text file consisting of a number of text records. One value, or data item, is separated from the next by a comma. The item can also be enclosed in double quotes - this is optional unless the text value contains a comma in which case it **must be** enclosed in double quotes. An empty text value is recorded as either "" or simply left blank (see examples of this below). The final item in each record does not have a comma after it.

For this census the file must contain a number of different record types. The type of record is identified by the first text value. There should be 1 header record and a minimum of 1 pupil record.

Please note: we have assumed that data supplied by local authorities will never contain double quotes (ie "). If this is not the case for you, can you please contact the department as special handling of such characters are required.

Be careful if using Microsoft Excel as the format of some fields may be changed, such as dates, when converting between Excel and CSV formats.

## 4.9 CSV header

The first record in any CSV file **must be** a header record. There can only be one header record in a file. A header record has the following structure:

Field	Description	Used in example
RecordType	Identifies the record type, must be "H".	H
Collection	Name of the data collection. Fixed value of 'APC'.	APC
Term	'SPR' denotes collection is carried out in the spring term.	SPR
Year	The year of the collection.	2017
Reference date	The reference date is normally the day of the census or collection and is used as the baseline date from which comparisons with other dates in	2017-01-19



Field	Description	Used in example
	the return can be made.	
Source level	Should be set to L for a local authority system.	L
LEA	Standard DfE three digit local authority number.	999
Software code	Software code is available for suppliers to provide their own identifier. Description: Software supplier reference.	Software supplier reference
Release	A code / date (or combination) provided by the software supplier to assist in identification.	
Serial number	An incremented number (starting at '001') generated by a provider's software, enables data collection systems to identify re-submissions.	001
Date & time	Date and time of generation of the return in CCYY-MM-DDThh:mm:ss format.	2017-01-19T10:36:42
CBDS level	Fixed value of 'pupil' shows which of the defined CBDS levels are present within a data return.	Pupil

Example:

H,APC,SPR,2017,2017-01-19 ,L,999,Software supplier reference,,001,2017-01-19T10:36:42,Pupil

If the export facility on a source system does not allow you to add a header record you will have to add it manually. For example, use a text editor (such as windows notepad) to open the export file and type in the appropriate header record at the beginning of the file. **Be careful not to accidentally change anything else within the file.**

## 4.10 CSV pupil record structure (single pupil record)

A Pupil record has the following structure:

Field	Description	Used in example
RecordType	Identifies the record type, must be "P".	P
UPN (100001)	A unique pupil number (UPN) is allocated to each pupil in maintained schools in England and Wales. It is an identifier for use in the educational context during a child's school career only and subject to data protection restrictions.	X123456789876
ULN (100016)	Unique learner number	1234567890
Surname (100003)	Full legal surname (derived from family, clan or marital association) of the child.	Familyname
Forename (100004)	Full given first name of the child (not common contractions).	Firstname
Middle names (100006)	In full, not shortened or familiar versions. If pupil has no middle name(s) then this field must be left blank.	Middlename
Pupil/child preferred surname (100011)	The surname preferred by the pupil (as written) most commonly used in the establishment (for pupils in alternative provision, the surname most commonly used in the LA).	Familyname
Date of birth (100292)	Date of birth.	1066-03-24
Gender (100287)	Gender of child (as defined in common data elements).	2
Child ethnicity (100563)	Child ethnicity code.	WBRI
Free school meal eligibility start date (100484)	Start date for pupil's free school meal eligibility.	1087-09-01
Free school meal eligibility end date (100485)	End date for pupil's free school meal eligibility.	1087-12-01
Country of UK (200634)	Country of United Kingdom in which the establishment is located.	ENG
Early years pupil premium eligibility (100559)	Records whether a child is eligible to receive early years pupil premium	Y

<b>Field</b>	<b>Description</b>	<b>Used in example</b>
Early years pupil premium basis funding (100560)	Records the basis under which early years pupil premium is funded	EE
Service children in education indicator (100330)	Indicates if a child has parent(s) who are service personnel serving in regular military units of all forces and exercising parental care and responsibility.	N
Post looked after arrangements (100587)	Indicates that a child has been looked after	N
Type of alternative provision (100471)	The type of provision, for example independent school, hospital, etc.	NOT
SEN provision (100580)	Provision types under the SEN code of practice.	A
Post code (100121)	The code allocated by the post office to identify a group of postal delivery points.	ZZ99 9XX
Unique property reference number (100586)	The unique property reference number	123456789102
Sub-dwelling (100103)	Flat, apartment name or number or other sub-division of a dwelling.	2 Flat
Dwelling (100109)	Dwelling name and/or number.	123
Street (100115)	Street name or street description that has been allocated to a street by the street naming authority.	Street Lane
Locality (100116)	The locality name refers to a neighbourhood, suburb, district, village, estate, settlement, or parish that may form part of a town, or stand in its own right within the context of an administrative area. Where an industrial estate contains streets it is defined as a locality in its own right.	Localarea
Town (100117)	The town name refers to a city or town that is not an administrative area, a suburb of an administrative area that does not form part of another town or a London district.	Townbury

Field	Description	Used in example
Administrative area (100118)	The administrative area is a geographic area that may be the highest level local administrative area, which may be a county or a unitary authority, an island or island group or London.	Postaltown
Post Town (100119)	The post office usually assigns these based on sorting office.	Countyshire
Address line 1 (100128)	First line of address.	
Address line 2 (100129)	Second line of address.	
Address line 3 (100130)	Third line of address.	
Address line 4 (100131)	Fourth line of address.	
Address line 5 (100132)	Fifth line of address.	

Example (this example uses the BS7666 (v1.4) format and leaves the Line Address items blank):

P,X123456789876,1234567890,Familyname,Firstname,Middlename,Familyname,1066-03-24,2,WBRI,1087-09-01,1087-12-01,ENG,Y,EE,N,N,NOT,A,ZZ99 9XX,123456789102, 2 Flat,123,Street Lane,Localarea,Townbury,Postaltown,Shireborough,,,,,

## 4.11 CSV pupil record structure (two pupil records)

Example of a complete return, with two pupils (All addresses are in Line Address format):

Field	Description	Used in example
RecordType	Identifies the record type, must be "H".	H
Collection	Name of the data collection. Fixed value of 'APC'.	APC
Term	'SPR' denotes collection is carried out in the spring term.	SPR
Year	The year of the collection.	2017

Field	Description	Used in example
Reference date	The reference date is normally the day of the census or collection and is used as the baseline date from which comparisons with other dates in the return can be made.	2017-01-19
Source level	Should be set to L for a local authority system.	L
LEA	Standard DfE three digit local authority number.	999
Software code	Software code is available for suppliers to provide their own identifier.  Description: Software supplier reference.	Software supplier reference
Release	A code / date (or combination) provided by the software supplier to assist in identification.	
Serial number	An incremented number (starting at '001') generated by a provider's software, enables data collection systems to identify re-submissions.	001
Date & time	Date and time of generation of the return in CCYY-MM-DDThh:mm:ss format.	2017-01-19T10:36:42
CBDS level	Fixed value of 'pupil' shows which of the defined CBDS levels are present within a data return.	Pupil
RecordType	Identifies the record type, must be "P".	P
UPN (100001)	A unique pupil number (UPN) is allocated to each pupil in maintained schools in England and Wales. It is an identifier for use in the educational context during a child's school career only and subject to data protection restrictions.	X123456789876
ULN (100016)	Unique learner number	1234567890
Surname (100003)	Full legal surname (derived from family, clan or marital association) of the child.	Familyname

<b>Field</b>	<b>Description</b>	<b>Used in example</b>
Forename (100004)	Full given first name of the child (not common contractions).	Firstname
Middle names (100006)	In full, not shortened or familiar versions. If pupil has no middle name(s) then this field must be left blank.	Middlename
Pupil/child preferred surname (100011)	The surname preferred by the pupil (as written) most commonly used in the establishment (for pupils in alternative provision, the surname most commonly used in the LA).	Familyname
Date of birth (100292)	Date of birth of the pupil.	1066-03-24
Gender (100287)	Gender of child (as defined in common data elements).	2
Child ethnicity (100563)	Child ethnicity code.	WBRI
Free school meal eligibility start date (100484)	Start date for pupil's free school meal eligibility.	1087-09-01
Free school meal eligibility end date (100485)	End date for pupil's free school meal eligibility.	1087-12-01
Country of UK (200634)	Country of United Kingdom in which the establishment is located.	ENG
Early years pupil premium eligibility (100559)	Records whether a child is eligible to receive early years pupil premium	Y
Early years pupil premium basis funding (100560)	Records the basis under which early years pupil premium is funded	EE
Service children in education indicator (100330)	Indicates if a child has parent(s) who are service personnel serving in regular military units of all forces and exercising parental care and responsibility.	N
Post looked after arrangements (100587)	Indicates that a child has been looked after	N
Type of alternative provision (100471)	The type of provision, for example independent school, hospital, etc.	NOT
SEN provision (100580)	Provision types under the SEN code of practice.	A

<b>Field</b>	<b>Description</b>	<b>Used in example</b>
Post code (100121)	The code allocated by the post office to identify a group of postal delivery points.	ZZ99 9XX
Unique property reference number (100586)	The unique property reference number	123456789102
Sub-dwelling (100103)	Flat, apartment name or number or other sub-division of a dwelling.	
Dwelling (100109)	Dwelling name and/or number.	
Street (100115)	Street name or street description that has been allocated to a street by the street naming authority.	
Locality (100116)	The locality name refers to a neighbourhood, suburb, district, village, estate, settlement, or parish that may form part of a town, or stand in its own right within the context of an administrative area. Where an industrial estate contains streets it is defined as a locality in its own right.	
Town (100117)	The town name refers to a city or town that is not an administrative area, a suburb of an administrative area that does not form part of another town or a London district.	
Administrative area (100118)	The administrative area is a geographic area that may be the highest level local administrative area, which may be a county or a unitary authority, an island or island group or London.	
Post town (100119)	The post office usually assigns these based on sorting office.	
Address line 1 (100128)	First line of address.	123 Street Lane
Address line 2 (100129)	Second line of address.	Localarea
Address line 3 (100130)	Third line of address.	Townbury
Address line 4 (100131)	Fourth line of address.	Postaltown
Address line 5 (100132)	Fifth line of address.	Countyshire
RecordType	Identifies the record type, must be "P".	P
UPN (100001)	As above in first pupil record.	Z987654321234
Surname (100003)	As above in first pupil record.	FamilynameX

Field	Description	Used in example
Forename (100004)	As above in first pupil record.	FirstnameX
Middle names (100006)	As above in first pupil record.	
Pupil/child Preferred Surname (100011)	As above in first pupil record.	
Date of birth (100292)	As above in first pupil record.	1067-02-02
Gender (100287)	As above in first pupil record.	1
Pupil ethnicity (100563)	As above in first pupil record.	MWBC
Free school meal eligibility start date (100484)	As above in first pupil record.	1087-09-01
Free school meal eligibility end date (100485)	Still eligible for free school meals.	
Country of UK (200634)	As above in first pupil record.	ENG
Early years pupil premium eligibility (100559)	Records whether a child is eligible to receive early years pupil premium	Y
Early years pupil premium basis funding (100560)	Records the basis under which early years pupil premium is funded	EO
Service children in education indicator (100330)	As above in first pupil record.	N
Post looked after arrangements (100587)	Indicates that a child has been looked after	N
Type of alternative provision (100471)	As above in first pupil record.	NOT
SEN provision (100580)	As above in first pupil record.	N
Post code (100121)	As above in first pupil record.	ZZ99 9YY
Unique property reference number (100586)	The unique property reference number	223456789102
Sub-dwelling (100103)	As above in first pupil record.	
Dwelling (100109)	As above in first pupil record.	
Street (100115)	As above in first pupil record.	
Locality (100116)	As above in first pupil record.	
Town (100117)	As above in first pupil record.	
Administrative area (100118)	As above in first pupil record.	



Field	Description	Used in example
Post town (100119)	As above in first pupil record.	
Address line 1 (100128)	As above in first pupil record.	321 Lane Road
Address line 2 (100129)	As above in first pupil record.	Localarea
Address line 3 (100130)	As above in first pupil record.	Countyshire
Address line 4 (100131)	As above in first pupil record.	
Address line 5 (100132)	As above in first pupil record.	

H,APC,SPR,2017,2017-01-19,L,999,Software supplier reference,,001,2017-01-19T10:36:42,Pupil

P,X123456789876,1234567890,Familyname,Firstmane,Middlename,Firstname,1066-03-24,2,WBRI,1087-09-01,1087-12-01,ENG,Y,EE,N,N,NOT,A,ZZ99 9XX,123456789102,,,,,,,,, "123, Street Lane",Localarea,Townbury,Countyshire,,

P,Z987654321234, , FamilynameX,FirstnameX ,,1067-02-02,1,MWBC,1087-09-01,,ENG,Y,EO,N,N,NOT,N,ZZ99 9YY, 223456789102,,,,,,,,,321 Lane Road,Localarea,Countyshire,,

## 4.12 Special notes for CSV returns

### File naming conventions

The department recommends that the file is given a meaningful name that includes the local authority number (the 3 digit serial number). The file extension should be .csv. This will help users in selecting the correct file to upload to COLLECT.

## 5 Validation checks

An Excel workbook specifies validation rules that are applied to the data by DfE using the COLLECT system. They should also be applied, as far as possible, by local authorities prior to loading the data onto the COLLECT system and any issues resolved.

Those data items included in the tables but not appearing in the validation rules - that is, those not shown in the validation rules as "must be present" - should be supplied where available (eg pupil's middle name).

The workbook also shows which validations are 'errors' and which ones are 'queries':

- an 'error' is a failed validation check that **must be corrected** before the return is submitted. There may be rare scenarios where a few 'errors' can be incorrectly triggered. The department will assess these scenarios on a case by case basis and, where appropriate, allow the 'error' to be 'OKable' in COLLECT
- a 'query' is one which **must be investigated**, and potentially corrected, as it identifies data that would usually be invalid. 'Queries' may be accepted with appropriate notepad entries. The department will accept notes at return level



Department  
for Education

© Crown copyright 2017

This publication (not including logos) is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

To view this licence:

visit [www.nationalarchives.gov.uk/doc/open-government-licence/version/3](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3)

email [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk)

write to Information Policy Team, The National Archives, Kew, London, TW9 4DU

About this publication:

enquiries <https://www.education.gov.uk/form/data-collection-request-form>

download [www.gov.uk/government/publications](http://www.gov.uk/government/publications)

Reference: DFE-00138-2016



Follow us on  
Twitter: [@educationgovuk](https://twitter.com/educationgovuk)



Like us on Facebook:  
[facebook.com/educationgovuk](https://facebook.com/educationgovuk)