



Information for Bulk Users

July 13

Contents

Introduction	2
File Specifications	2
Useful links.....	5

Information for Bulk Users

Introduction

This information explains the format, naming and numbering of the data files that are produced from the bulk data. It also includes some useful links for further information from the census section of the Office for National Statistics (ONS) website.

File types

The Output Files

All the output files are in a comma separated format (.CSV). The census bulk format contains 4 file types (N.B. an extract may produce more than 4 files as described in the CSV format section):

- Meta - Table description file. This file contains text that describes the table.
- Desc - Textual description file. This file contains the textual descriptions of the cells.
- Data - Data file. This file holds the data values in row major order.
- Code - Textual description file. This file includes code information.

The CSV Format

- Each new line in the CSV file uses the carriage return character (ASC11 code 13) followed by the line feed character (ASC11 code 10).
- Each line in a given CSV file contains the same number of fields separated by a comma.
- All text fields are delimited using double quotes.
- There is a maximum number of 250 fields per CSV file. When a census extract file requires more than 250 fields the naming convention for multiple files described in this document is adopted.

File Naming

All files produced as a result of the extract process carry the extension: .CSV to indicate its contents are in comma separated form as described in this document.

Table type

The first two characters indicate the table type:

QS	Quick Statistics
KS	Key Statistics
DC	Detailed Characteristics
CT	Commissioned Tables
LC	Local Characteristics

Table number

The following four characters are digits; they indicate the table number.

Country

The following 2 characters indicate the country as follows:

EW	England & Wales
WA	Wales

Geography Labels

r	Region
la	Merged Wards

File type

The next four characters will indicate the type of file:

DESC	Table Description File
CODE	Table Desc File
META	Textual Description File
DATA	Data File

Split file

If an extract file contains more than 250 fields then the file will be split. The next character indicates if the file is split and if so gives its sequence position:

0	No split
A	1 st file in split
B	2 nd file in split
C	3 rd file in split
D	... and so on

Geography level (DATA files)

The next character applies to DATA files only and is used to indicate the geography level of the counts it contains. Areas contained within each level of geography are identified as follows:

OA (where applicable)

1. Highest geography level, England and Wales combined.
2. England, Wales.
3. Regions and Wales.
4. Unitary Authorities, Non Metropolitan Districts, Metropolitan Districts, London Boroughs, Welsh Unitary Authorities.
5. Super Output Areas Middle Layers, England and Wales.
6. Super Output Areas Lower Level, England and Wales.

Please note: CSV files that end with _A etc, contain Output Areas by Region

WARDS (where applicable)

1. Highest geography level, England and Wales combined.
2. England, Wales.
3. Regions and Wales.
4. Unitary Authorities, Counties, Metropolitan Counties, Inner and Outer London, Welsh Unitary Authorities.
5. Unitary Authorities, Non Metropolitan District, Metropolitan Districts, London Boroughs, Welsh Unitary Authorities.
6. Electoral Wards /Divisions and Welsh Divisions

Example file names

The example file name **DC6201EWDData V3 CSV** would therefore hold all the values for Detailed Characteristics table 6201 for England and Wales at Region and Wales level.

Useful links

Census landing page

For further information on the glossary, definitions, and census prospectus

<http://www.ons.gov.uk/ons/guide-method/census/2011/index.html>

ONS Geography

For further information about census geographies

<http://www.ons.gov.uk/ons/guide-method/geography/ons-geography/index.html>

GIS boundaries for use with statistics

Download digital boundaries for 2011 output areas, super output areas (LSOAs and MSOAs) and workplace zones for England and Wales. These boundaries are for users to input into their own geographical information systems to carry out spatial analysis or visualisation.

<http://www.ons.gov.uk/ons/guide-method/geography/products/census/spatial/2011/index.html>

Census geography look-ups

ONS Geography provides a range of products to define names and codes and the relationships between geographies. These products give information on the relationship between 2011 output areas and other geographies and the changes between 2001 and 2011 output areas.

<http://www.ons.gov.uk/ons/guide-method/geography/products/census/lookup/index.html>