

## Error in the distribution of Religion categories in Camden, Islington and Tower Hamlets

June 2015

### The error

On 26 February 2015, ONS announced that an error had been discovered in the processing of counts of religion for three London boroughs: Camden, Islington and Tower Hamlets.

'Religion not stated' had been overestimated by 61,500 persons across the three boroughs, with all other categories underestimated by same number. Estimates for all other local authorities in England and Wales were unaffected. 'Religion not stated' and other categories of religion counts for Inner London, London and national totals were consequently also incorrect by the same number as the as the three boroughs combined.

### ONS information on error

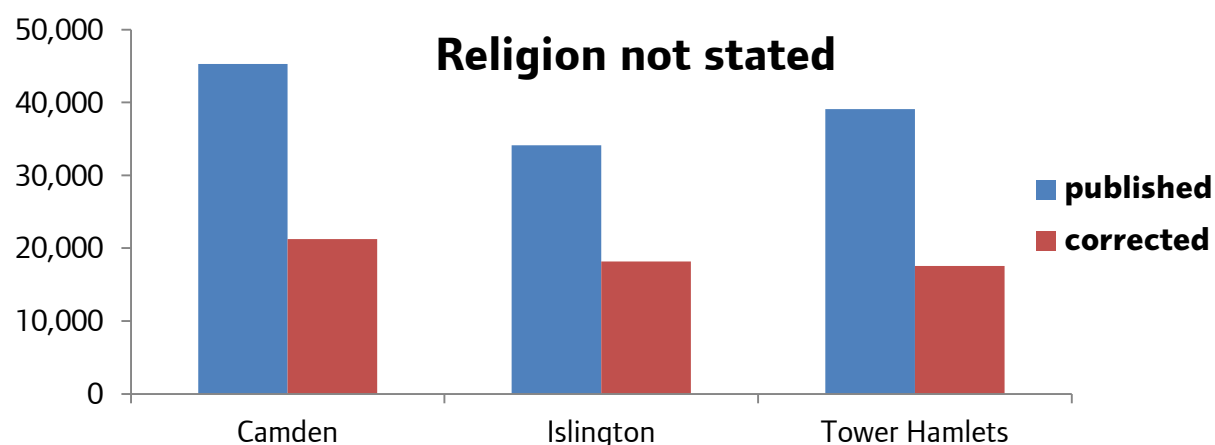
ONS have said that they will not correct the census database for the error, nor will they publish any corrected tables. They have released estimates of corrected borough usual residents by detailed age and sex, equivalent to Detailed Characteristics table DC2107EW, although some small cell counts in these tables have not been independently adjusted. Figures for males, females and persons have been estimated separately with the sum of males and females not equal to persons.

Because of the limitations in data supplied by ONS, GLA have produced estimates of corrected figures for wards and MSOAs in the three boroughs. This note discusses the size and incidence of the error and describes the methodology used in calculating the corrected tables.

## The size of the error

Figure 1 below shows the size of the error in the 'Religion not stated' category for the three boroughs. Camden had the biggest correction 24,000 some 53 per cent of the published figure; Tower Hamlets' correction was 21,500 some 55 per cent, and Islington was down 15,900 or 47 per cent.

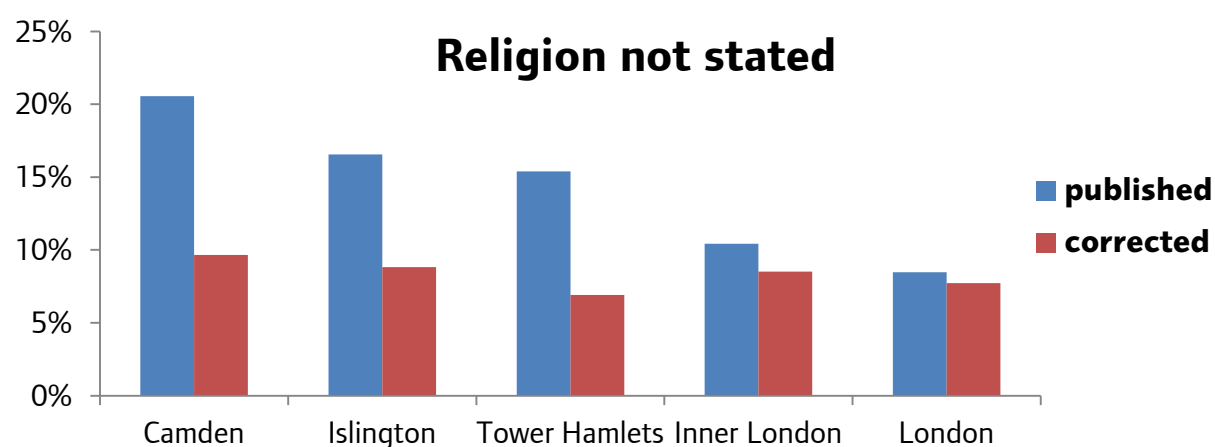
**Figure 1: Correction to 'Religion not stated' in the three boroughs**



The correction also had a significant effect on Inner London and all-London figures as shown in figure 2. The proportions 'Religion not stated' fell from 10.4 per cent to 8.5 per cent in Inner London and 8.5 per cent to 7.7 per cent for London.

For England and Wales, 'Religion not stated' fell from 7.2 per cent to 7.1 per cent.

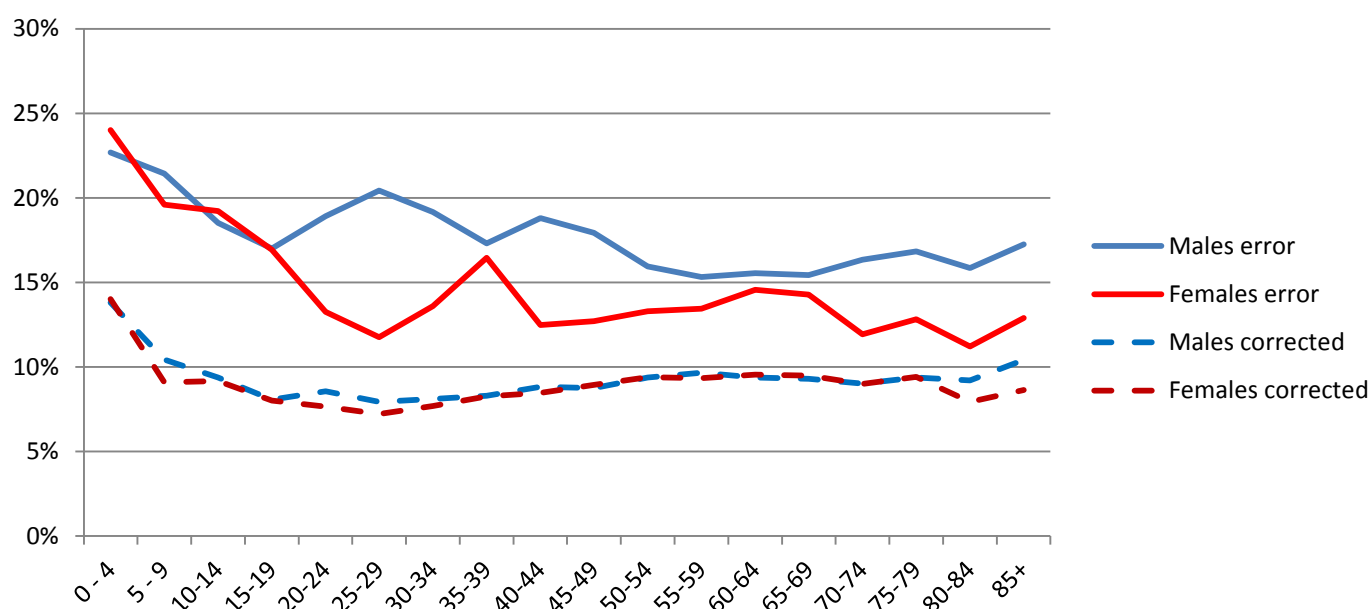
**Figure 2: Percentage correction to 'Religion not stated' in the three boroughs and London**



## The incidence of the error by sex and age

Unlike other areas, published figures for Camden, Islington and Tower Hamlets showed considerably higher Religion not stated for males than females, and showed a peak amongst persons aged 20 to 39. Corrected figures from ONS suggest these features were largely an artefact of the error. Figure 3 shows the 'Religion not stated' rates by age and sex for Islington both published and corrected. Published 'Religion not stated' rates were consistently higher for males, but the corrected rates are nearly identical. The same pattern is seen for Camden and Tower Hamlets. Figure 3 also shows that children aged 0 to 4 had the highest rates of 'Religion not stated' of any age group both before and after the correction. This was also true for other local authority areas.

**Figure 3: 'Religion not stated' by age and sex for Islington**



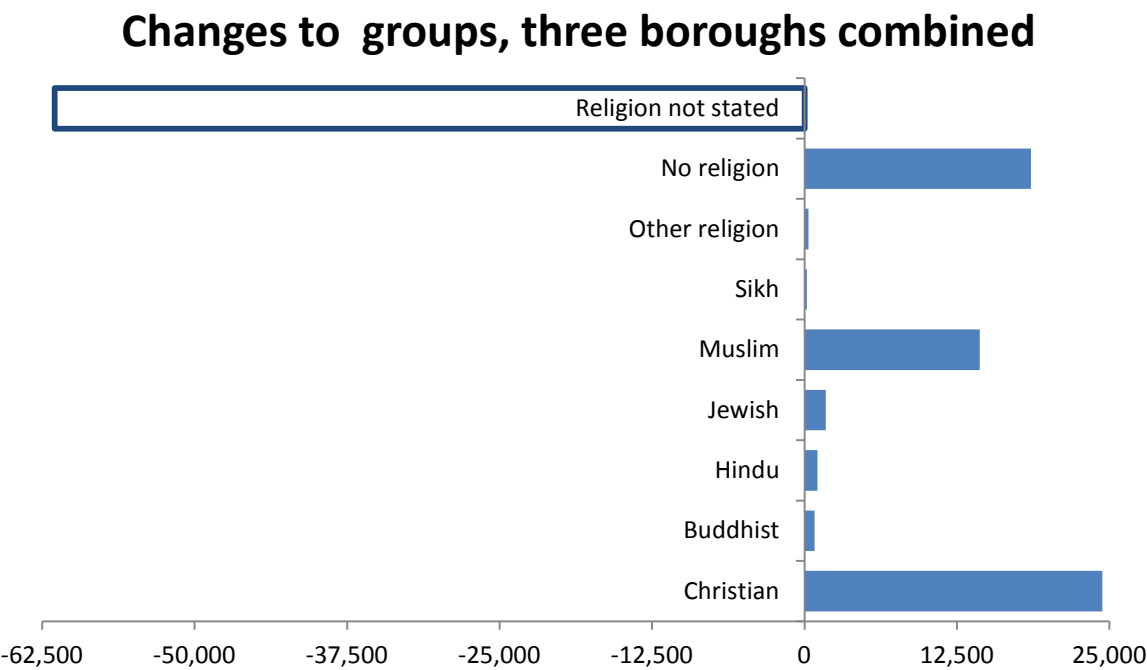
# Effect of error on other religions

Since the error misallocated persons in other groups to ‘Religion not stated’, correcting the error increases the numbers in other categories, that is all religions and No religion. For each borough, all these groups increase by about the same proportions – 13 per cent for Camden, 9 per cent for Islington and 10 per cent for Tower Hamlets.

These corrections represented much smaller proportionate changes to Inner London and all London figures. Muslims, the religion most concentrated in the three boroughs compared to the rest of the capital increased by 3 per cent in Inner London, and across London by 1.4 per cent. Hindus and Sikhs, both uncommon religions in the three boroughs had the smallest increases: Counts of Hindus increased 1.5 per cent in Inner London and 0.3 per cent in London; Sikhs, concentrated strongly in Outer London, increased by 1.3 in Inner London and only 0.1 per cent in London.

Figure 4 shows the changes to the three boroughs combined. Christian, the largest group increased by 24,400, No religion by 18,600 and Muslim by 14,400. All other groups combined increased by 4,100. Table 1 on the next page shows the main groups for the boroughs separately.

Figure 4: Changes to religion categories for three boroughs



**Table 1: Published and corrected figures for three boroughs**

	Christian	Muslim	All other religions	No religion	Not stated
PUBLISHED					
Camden	74,800	26,600	17,500	56,100	45,300
Islington	82,800	19,500	7,700	61,900	34,100
Tower Hamlets	68,800	87,700	9,900	48,600	39,100
Percentages					
Camden	34.0	12.1	7.9	25.5	20.5
Islington	40.2	9.5	3.7	30.0	16.6
Tower Hamlets	27.1	34.5	3.9	19.1	15.4
CORRECTED					
Camden	84,900	30,400	19,900	63,900	21,300
Islington	90,400	21,300	8,400	67,800	18,200
Tower Hamlets	75,600	96,500	10,800	53,600	17,600
Percentages					
Camden	38.5	13.8	9.0	29.0	9.6
Islington	43.9	10.3	4.1	32.9	8.8
Tower Hamlets	29.7	38.0	4.3	21.1	6.9

‘All other religions’ comprises Buddhist, Hindu, Jewish, Sikh and ‘Other religions’.

## The need for a set of corrected tables

ONS have advised users to correct data for Camden, Islington and Tower Hamlets by applying correction factors derived from their age by sex estimates for the three boroughs. There are several problems with this:

- Factors have been estimated individually so will not produce tables that are internally consistent: estimates of individual ages will not generally add to all ages, and males plus females persons will not add to persons.
- Totals between tables will differ for the same reason and there will be inconsistency at different geographies.
- There are no factors and so no method for estimating tables where the population is other than usual residents, where the population’s religion distribution might be expected to differ from usual residents. Instances of this include: Household Representative Persons, resident population in work, residents in communal establishments.
- Corrected figures for Inner London, London and England and Wales will also be required for comparisons. ONS have published corrected figures for England and Wales, but not for the other aggregates. As stated above, figures for different geographies will not be consistent.

GLA undertook to produce a common set of corrected tables that would be available for all users, ameliorating where possible the issues above.

## GLA methodology for producing corrected tables

There is no other information available to correct tables so the basic method used is the same as that recommended by ONS: apply the borough level factors for the appropriate population for each table. We have however added several refinements:

To avoid inconsistency between rows, each table has only been estimated once using only the most detailed cells in the table, and all other cells calculated by summation.

To avoid inconsistency between columns, each table is further pro-rated to match marginal totals, and

Once borough tables for Camden, Islington and Tower Hamlets have been calculated, the figures from these are used both to estimate lower geographies (by simple application of the same factors), and calculate higher geographies (by summation).

### Iterative proportional fitting

The method used to ensure the corrected tables agree with marginal totals by religion and other variables is a version of the iterative proportional fitting procedure (or more succinctly, the RAS algorithm). 'RAS-ing' a table just means cells are alternately pro-rated to row and column totals. At each iteration of pro-rating, the differences between cell and marginal totals will decrease. The procedure continues until the differences are small enough to be disregarded.

In summary the method is:

1. Output uncorrected table without subtotals
2. Apply correction factors to most detailed sub-tables. The results will agree with the corrected religion totals but other marginal totals will not agree with un-corrected tables.
3. Apply iterative pro-rating to each sub-table.
4. Calculate final adjustment factors from pro-rated tables
5. Apply factors to lower geographies for small area estimates
6. Sum estimates to produce subtotals and totals and figures for Inner London, London etc.

Iterative proportional fitting has been applied individually to Camden, Islington and Tower Hamlets to ensure tables are consistent at borough and higher geographies. Unfortunately it does not ensure that the tables are consistent for wards (or other lower geographies).

## Calculating factors for populations other than usual residents

As stated previously, ONS's estimates of corrected figures are for usual residents only. Below we state the estimates used for other populations:

### HRP Tables

There is limited information on HRPs by age that could be used with figures of usual residents by age and religion to indirectly model the distribution of the religion error for HRPs. However initial investigations into this did not identify a consistent effect that could be applied to tables, and so this approach was not pursued.

Instead, the information used was aged implied by the type of household in the household composition table DC1202EW:

- One person households aged 65 and over; One family households, all aged 65 and over; and Other household types all aged 65 and over, HRP was necessarily aged 65 and over. (These three groups made up around three-quarters of all HRPs aged 65 and over.)
- One person households aged under 65; all household types with dependent children; and All full-time student households, HRP assumed to be aged 16-64, and
- All other household types, HRP any age 16 and over.

Correction factors for usual residents aged 16-64, 65 and over or all aged 16 and over were applied to table DC1202EW. Totals of HRPs by religion from DC1202EW were used as controls for the other HRP tables.

### Industry and Occupation Tables

The population for these tables is usual residents aged 16 and over in work. It might be possible to model the distribution the religion error using age-dependent characteristics all residents, but this was not pursued. The distribution of all usual residents aged 16 and over in work was calculated from Economic Activity table DC6205EW, so corrected totals from DC6205EW were taken as controls for industry and occupation tables.

### Communal Establishment Table

DC4409EWLA, Communal establishment type by religion by sex, was the only table of CE residents by Religion. As there is no age breakdown of Communal Establishment residents by religion, there is no information to estimate the distribution of errors by establishment type and we would have to apply factors based on all ages of residents across the CE types which is equivalent to the implausible assumption of no variation by age. This table was hence not corrected.

## Available data

Re-estimated data for the three boroughs and all other merged local authorities in England have been produced as SASPAC system files [on the SASPAC website](#)

There are three separate files, containing wards, MSOAs and local authority and higher geographies. Data for local authorities other than Camden, Islington and Tower Hamlets are the same as the uncorrected datasets.

## Worked example

DC3203EW for 16 to 49 year old males in Islington

Original uncorrected table

	All persons	Christian	Buddhist	Hindu	Jewish	Muslim	Sikh	Other religion	No religion	Religion not stated
Good/Very Good Health										
limited a lot	414	127	3	1	3	62	1	4	108	105
limited a little	1,197	369	15	10	14	126	5	9	381	268
not limited	55,852	17,534	497	704	585	4,016	151	248	21,850	10,267
Fair Health										
limited a lot	754	274	10	5	3	97	6	5	180	174
limited a little	1,355	457	12	13	6	172	3	18	397	277
not limited	2,516	710	46	17	10	314	6	17	847	549
Bad/Very bad Health										
limited a lot	1,516	532	15	7	5	233	8	18	370	328
limited a little	480	119	8	2	1	104	1	8	131	106
not limited	251	55	3	3	4	49	0	3	66	68

Corrected totals can be taken directly from ONS's spreadsheet. (For tables based on HRP's we have to calculate these totals)

sum from original table	64335.0	20177.0	609.0	762.0	631.0	5173.0	181.0	330.0	24330.0	12142.0
corrected figures	64335.0	22804.8	684.6	861.1	714.6	5822.5	205.1	372.5	27536.7	5333.1

Pro-rated to corrected religion – totals by health/disability do not match

	All persons	Christian	Buddhist	Hindu	Jewish	Muslim	Sikh	Other religion	No religion	Religion not stated
Good/Very Good Health										
limited a lot	395.2	143.5	3.4	1.1	3.4	69.8	1.1	4.5	122.2	46.1
limited a little	1167.6	417.1	16.9	11.3	15.9	141.8	5.7	10.2	431.2	117.7
not limited	56045.0	19817.6	558.7	795.6	662.5	4520.2	171.1	279.9	24729.8	4509.5
Fair Health										
limited a lot	731.7	309.7	11.2	5.7	3.4	109.2	6.8	5.6	203.7	76.4
limited a little	1339.8	516.5	13.5	14.7	6.8	193.6	3.4	20.3	449.3	121.7
not limited	2463.9	802.5	51.7	19.2	11.3	353.4	6.8	19.2	958.6	241.1
Bad/Very bad Health										
limited a lot	1486.2	601.3	16.9	7.9	5.7	262.3	9.1	20.3	418.8	144.1
limited a little	468.9	134.5	9.0	2.3	1.1	117.1	1.1	9.0	148.3	46.6
not limited	236.6	62.2	3.4	3.4	4.5	55.2	0.0	3.4	74.7	29.9

So pro rate to health/disability totals. Now religion totals won't match but they will be much closer than the original uncorrected table



	All persons	Christian	Buddhist	Hindu	Jewish	Muslim	Sikh	Other religion	No religion	Religion not stated
Good/Very Good Health										
limited a lot	414.0	150.4	3.5	1.2	3.6	73.1	1.2	4.7	128.0	48.3
limited a little	1197.0	427.5	17.3	11.6	16.3	145.4	5.8	10.4	442.1	120.7
not limited	55852.0	19749.4	556.7	792.8	660.3	4504.7	170.5	279.0	24644.6	4494.0
Fair Health										
limited a lot	754.0	319.1	11.6	5.8	3.5	112.5	7.0	5.8	209.9	78.7
limited a little	1355.0	522.4	13.6	14.9	6.9	195.8	3.4	20.5	454.4	123.0
not limited	2516.0	819.4	52.8	19.6	11.6	360.9	6.9	19.6	978.9	246.2
Bad/Very bad Health										
limited a lot	1516.0	613.3	17.2	8.1	5.8	267.5	9.2	20.7	427.2	147.0
limited a little	480.0	137.7	9.2	2.3	1.2	119.8	1.2	9.2	151.8	47.7
not limited	251.0	66.0	3.6	3.6	4.8	58.5	0.0	3.6	79.3	31.7
sums by religion		22805.2	685.6	859.9	713.8	5838.2	205.3	373.6	27516.2	5337.3

Repeat alternately pro-rating until desired precision is achieved. This table took three iterations, but some took a lot more.

	All persons	Christian	Buddhist	Hindu	Jewish	Muslim	Sikh	Other religion	No religion	Religion not stated
Good/Very Good Health										
limited a lot	414.0	150.4	3.5	1.2	3.6	72.9	1.2	4.7	128.2	48.3
limited a little	1197.0	427.6	17.3	11.6	16.3	145.0	5.8	10.4	442.5	120.6
not limited	55852.0	19748.2	555.9	793.9	661.1	4492.2	170.4	278.1	24662.1	4490.2
Fair Health										
limited a lot	754.0	319.2	11.6	5.8	3.5	112.2	7.0	5.8	210.1	78.7
limited a little	1355.0	522.5	13.6	14.9	6.9	195.3	3.4	20.5	454.9	123.0
not limited	2516.0	819.6	52.7	19.7	11.6	360.0	6.9	19.5	979.9	246.1
Bad/Very bad Health										
limited a lot	1516.0	613.6	17.2	8.1	5.8	266.9	9.2	20.7	427.7	146.9
limited a little	480.0	137.8	9.2	2.3	1.2	119.6	1.2	9.2	152.0	47.6
not limited	251.0	66.0	3.6	3.6	4.8	58.4	0.0	3.6	79.4	31.7
sums by religion		22804.8	684.6	861.1	714.6	5822.5	205.1	372.5	27536.7	5333.1
difference from target		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Divide by uncorrected figures to give final correction factors, One empty cell in this table set to 0:

	All persons	Christian	Buddhist	Hindu	Jewish	Muslim	Sikh	Other religion	No religion	Religion not stated
Good/Very Good Health										
limited a lot		1.1844	1.1762	1.1859	1.1883	1.1763	1.1864	1.1791	1.1869	0.4599
limited a little		1.1588	1.1508	1.1603	1.1627	1.1509	1.1608	1.1537	1.1613	0.4500
not limited		1.1263	1.1185	1.1278	1.1300	1.1186	1.1282	1.1213	1.1287	0.4373
Fair Health										
limited a lot		1.1650	1.1569	1.1665	1.1689	1.1570	1.1669	1.1598	1.1675	0.4524
limited a little		1.1433	1.1354	1.1448	1.1472	1.1355	1.1453	1.1383	1.1458	0.4440
not limited		1.1544	1.1464	1.1559	1.1582	1.1465	1.1563	1.1493	1.1569	0.4483
Bad/Very bad Health										
limited a lot		1.1534	1.1454	1.1549	1.1572	1.1455	1.1553	1.1482	1.1558	0.4479
limited a little		1.1576	1.1496	1.1591	1.1615	1.1497	1.1596	1.1525	1.1601	0.4495
not limited		1.1998	1.1915	1.2014	1.2038	1.1916	0.0000	1.1945	1.2024	0.4659

Then apply these factors to uncorrected table to get figures for boroughs and wards.

Merge corrected data for Camden, Islington and Tower Hamlets with data for rest of England, then sum for Inner London and higher aggregates.

Finally calculate totals by Health, Disability, All males and females and for persons.

Because of the RASing, the borough tables will be consistent. As stated before, the ward tables will not.

## Annex 1

### Tables that include a breakdown by Religion

Reference	population base	breakdown(s) in addition to religion
-----------	-----------------	--------------------------------------

#### Detailed Characteristics

DC1202EW	All Household Representative Persons	Household composition
DC2107EW	All usual residents	age by sex
DC2201EW	All usual residents	Ethnic group
DC2204EW	All usual residents	National Identity
DC2207EW	All usual residents	Country of birth by sex
DC3203EW	All usual residents	Broad age by sex by disability by health
DC4204EW	All Household Representative Persons	Tenure by car availability
DC4207EW	All Household Representative Persons	Occupancy rating (rooms) by Shared/unshared dwelling
DC4208EW	All Household Representative Persons	Occupancy rating (bedrooms) by Shared/unshared dwelling
DC4409EWLA *	All residents in Communal Establishments	Establishment type by sex
DC5204EW	All usual residents aged 16 and over	Qualifications by broad age
DC6205EW	All usual residents aged 16 and over	Economic activity by sex by broad age
DC6207EW	All usual residents aged 16 and over	NS-SEC by sex by broad age
DC6212EW	All usual residents aged 16+ in work	Industry by broad age
DC6214EW	All usual residents aged 16+ in work	Occupation by sex by broad age
DC6217EW	All usual residents aged 16+ in work	Industry by sex

\*Estimates have been calculated for all these DC tables except DC4409EWLA. There is no age breakdown of Communal Establishment residents by religion so we would have to apply factors based on all ages of residents which might not be appropriate.

Other tables including a breakdown by religion have not been re-estimated at this time by GLA. These tables comprise:

#### Key Statistics

KS209EW	All usual residents	religion only
---------	---------------------	---------------

#### Quick Statistics

QS208EW	All usual residents	religion only
QS210EW	All usual residents	religion with 'other religions' detailed

#### Workplace Statistics

WP210EW	All usual residents aged 16-74 in work in the area	religion only
---------	---	---------------

## Workday Statistics

WD210EW	All usual residents aged 16-74 in work in the area or not in work, resident in area	religion only
---------	--	---------------

## Workplace Statistics

WP210EW	All usual residents aged 16-74 in work in the area	religion only
---------	---	---------------

## Local Characteristics

LC1202EW	All Household Representative Persons	Household composition
LC2107EW	All usual residents	age by sex
LC2120EW	All usual residents in households	Accommodation type
LC2201EW	All usual residents	Ethnic group
LC2204EW	All usual residents	National Identity
LC2207EW	All usual residents	Country of birth by sex
LC3203EW	All usual residents	Broad age by sex by health
LC3207EW	All usual residents	Broad age by sex by disability
LC4204EW	All Household Representative Persons	Tenure by car availability
LC4207EW	All Household Representative Persons	Occupancy rating (rooms)
LC4208EW	All Household Representative Persons	Occupancy rating (bedrooms)
LC4417EW	All usual residents in households	Tenure by car availability
LC5204EW	All usual residents aged 16 and over	Qualifications
LC6205EW	All usual residents aged 16 and over	Economic activity
LC6207EW	All usual residents aged 16 and over	NS-SEC
LC6212EW	All usual residents aged 16+ in work	Broad Industry
LC6214EW	All usual residents aged 16+ in work	Occupation