

Technical Guidance on production of official statistics for Settlements and Urban-Rural Classification



(May 2016)

Introduction

1. In March 2015, the Northern Ireland Statistics and Research Agency (NISRA) published a [Review of the Statistical Classification and Delineation of Settlements](#). The work, which was commissioned by the Statistics Co-ordinating Group (SCG) Of the Northern Ireland Civil Service (NICS), was undertaken by a representative cross-Departmental sub-group and was based on three data sources, namely:
 - a. The 2011 Census of population;
 - b. November 2014 Settlement Development Limits provided by the Planning Section of the Department of the Environment; and
 - c. The drive times to town centres within settlements of 10,000 people or more, provided by the Analytical Service Unit of the Department of Social Development.This report replaced the [2005 Review](#), which was based on the 2001 Census and the Settlement Development Limits at the time.
2. Settlements can be grouped with other settlements of similar characteristics. The report used 8 Bands (A-H) based on the 2011 Census population only, and a default classification where settlements in Band A-E (i.e. those with a population greater than or equal to 5,000 people) are classified as urban, and Bands F-H (i.e. those with a population of less than 5,000 people) are classified as rural. The report also provides information on travel times to large settlements (i.e. settlements with a population of 10,000 people or more), which allows users to create their own grouping of settlements based on a combination of population size and drive-times.
3. This guidance follows the same structure as similar guidance produced for creating statistics for the 11 new [Local Government Districts](#) and the 80 new [District Electoral Areas](#).
4. The purpose of this paper is to provide further guidance on the production of official statistics for settlements or groups of settlements. In particular, how to take aggregate official spatial statistics and create statistics for these sub-geographies. It is worth noting that some issues associated with producing statistics for settlements become more pertinent when looking at smaller settlements.

Official Statistics

5. Currently, producers can create statistics for the new settlements from existing microdata (or unit record data) in a number of ways. These include:
- using the exact location (grid reference) of the household or business to spatially determine the appropriate settlement; or
 - using the postcode of the household or business to spatially determine the appropriate settlement; or
 - obtaining the settlement of the household or business directly without access to the address or postcode.

If access to existing microdata is available, statistics should be created in that order to ensure the highest level of accuracy.

6. The 2015 Settlements were first included in the November 2015 Central Postcode Directory (CPD). For each postcode, the CPD provides the following information:
- a. Settlement name and code;
 - b. Settlement band (A-H) – based on the 2011 Census population;
 - c. Default urban-rural classification;
 - d. Whether it lies within a 20 minute drive time of a town centre in a settlement with a population of 10,000 people or more (i.e. bands A-D);
 - e. Whether it lies within a 30 minute drive time of a town centre in a settlement with a population of 10,000 people or more (i.e. bands A-D);
 - f. Banded drive time to Belfast City Centre; and
 - g. Banded drive time to Derry City Centre.
7. NISRA have published a [methodology/quality paper](#) on creating a postcode-based lookup table to different geographies, including some quantification of accuracy levels. In short, postcode-based statistics are within 1.5 per cent of actual figures for Local Government Districts, Parliamentary Constituencies and District Electoral areas. For smaller geographies such as Wards and Super Output Areas, this percentage rises to 3-5 per cent, and even higher for 2011 Census Small Areas.
8. Full details of this exercise are reported in [CPD Accuracy](#), including the new settlements. In general, the level of accuracy decreases for smaller settlements as there will be more postcodes with associated properties that straddle the settlement boundaries relative to the number of postcodes where all properties are within the settlement boundary.

New Lookup Tables

9. For those without access to record-level address information, one-to-one relationship lookup tables were developed to create statistics for settlements. Separate lookup tables are created using Small Areas, Super Output Areas and the former Electoral Wards respectively as the building blocks. Whilst the approximation of settlements is better using the smallest building blocks (i.e. Small Areas), it is recognised that not all statistics are available at that level. For example, at the time of writing, of the 3,168 [datasets](#) available on the Northern Ireland Neighbourhood Information Service ([NINIS](#)) website, there were 323 datasets for Small Areas, 536 datasets for Super Output Areas, and 913 datasets for the former electoral wards. This compares to 50 datasets for the 2015 settlement classification. **The new lookup tables are available [here](#).**

Small Area to Settlement lookup table

10. The methodology used to create the Small Area level lookup table was based on the 2011 Census population, which was also used to group settlements by size. In the first step, a Small Area was allocated to a particular settlement if its boundary fell within that settlement in its entirety. This was the case for 2,917 of the 4,537 Small Areas (64.3 per cent). In addition, 877 Small Areas (19.3 per cent) did not have an overlapping land area with any settlement with more than 1,000 people, and thus were classified as Band H (open countryside). For the purpose of this lookup table, small villages with a population less than 1,000 people were added to this generic group.
11. In the second step, the remaining 743 Small Areas were allocated to the settlement or open countryside (Band H) that contained the majority (i.e. 50 % or more) of the Small Area population. For the vast majority of these cases, the Small Area straddled the settlement boundary, i.e. one part of the Small Area was within a settlement boundary, whilst the remainder of the Small Area was in open countryside.
12. Only seven Small Areas did not have an outright majority of the population in a particular settlement or open countryside. This is out of the 22 Small Areas which overlapped with two different settlements and open countryside. For these Small Areas, it was first considered whether the majority of the population resides in an urban settlement. If this was the case, the Small Area was allocated to the urban settlement with the largest population share. Alternatively, the Small Area was allocated to the rural settlement or rural hinterland with the largest population share.

13. Annex A provides the 2011 Census population estimates by Settlement Band, comparing counts based on exact boundaries with those that used grid squares, postcodes, and Small Areas as building blocks for approximation. Clearly, the smallest building blocks result in the best approximation, and conversely, approximations of larger settlements are better than those of smaller settlements.

Small Area to Urban-Rural Classification

14. Although Small Areas can be allocated to individual settlements and thus settlement bands and the default urban-rural split, it was found that a number of these areas could be better described as '*mixed urban-rural*'. In these Small Areas, the proportion of the population within an urban settlement (Band A-E) lies between 10 and 90 per cent. This cut-off was selected as there was a uniform distribution of urban proportions for all Small Areas within this range.

15. Table 1 shows for each settlement band the number of Small Areas that were either (1) entirely within a settlement or open countryside, (2) had a majority of 90 per cent or more of its population within the settlement boundary, or (3) had a majority of less than 90 per cent of its population within the settlement boundary. The latter group, accounting for 3.6 per cent of all Small Areas, are classified as '*mixed urban-rural*'.

Table 1: Allocations of Small Areas to Settlement Bands

Settlement Band	Number of Small Areas where either the entire or the majority of the population falls within the settlement boundary				Proportion of Small Areas where entire population falls within the settlement boundary
	Entire	90 per cent or more	Less than 90 per cent ('mixed')	Total	
Band A	800	10	6	816	98.0%
Band B	177	11	6	194	91.2%
Band C	1,273	97	44	1,414	90.0%
Band D	255	50	18	323	78.9%
Band E	219	49	37	305	71.8%
Band F	94	92	0	186	50.5%
Band G	99	147	3	249	39.8%
Band H	877	122	51	1,050	83.5%
TOTAL	3,794	578	165	4,537	83.6%

16. As settlements become smaller, the proportion of Small Areas that fully nest within a settlement boundary decreases. With this, the level of accuracy associated with approximating settlements by groups of Small Areas also tends to decrease. The lookup tables are accompanied by quality information highlighting how well the 2011 Census population and households estimates for each settlement can be approximated by groups of Small Areas.

Super Output Area to Urban-Rural Classification

17. The lookup table for Super Output Areas (SOAs) only includes the default classification whereby SOAs are regarded as urban, rural or mixed urban-rural, on the basis of the approach similar to that used for Small Areas. As Super Output Areas are built up from on average nine Small Areas, there is a greater likelihood that an SOA will contain parts of different settlements, thus making it less likely that more than 90 per cent of the population lies within one particular settlement. Indeed, 8.2 per cent of Super Output Areas are classified as '*mixed urban-rural*' compared to 3.6 per cent of Small Areas.

18. There were 396 SOAs (44.5 per cent) where the entire population was within a settlement with a population of 5,000 people or more (default urban area). Similarly, there were 231 SOAs (26.0 per cent) where the entire population was within the open countryside or a settlement with less than 5,000 people (default rural area). The remaining 263 SOAs (29.6 per cent) were classified as urban, rural or mixed depending on the proportion of the population within an urban settlement. Super Output Areas where this proportion lay between 10 and 90 per cent are classified as '*mixed urban-rural*'. Table 2 summarises this classification.

Table 2: Default urban-rural classification of Super Output Areas

Default Urban-Rural Classification	Number of SOAs	Proportion of SOAs
Entire population in urban settlement	396	44.5%
90 per cent or more of the population in urban settlements	155	17.4%
Total urban Super Output Areas	550	61.8%
Entire population outside urban settlements	231	26.0%
90 per cent or more of the population outside urban settlements	36	4.0%
Total rural Super Output Areas	267	30.0%
Mixed urban-rural Super Output Areas	73	8.2%
ALL Super Output Areas	890	100.0%

Electoral Ward to Urban-Rural Classification

19. Wards are generally too large to classify them as either urban or rural. In the 2005 classification, '[mixed urban-rural](#)' was introduced for Wards that consist of both urban and rural 2001 Census Output Areas. To remain consistent with the approach taken for Small Areas and Super Output Areas, Wards where the proportion of the population in urban settlements lay between 10 and 90 per cent were classified as 'mixed urban-rural'. This led to 314 urban (54.0 per cent), 212 rural (36.4 per cent) and 56 mixed urban rural Wards (9.6 per cent). The lookup table also includes for each Ward the number of Small Areas that are classified as urban, rural or mixed urban-rural.

20. Annex B provides 2011 Census counts of both the total population and total number of households for urban, rural and mixed urban-rural areas, using Small Areas, Super Output Areas and Wards as building blocks for approximation. It is evident that the selection of the smallest building block returns the best approximation to the published counts, which are based on exact boundaries. This is a direct result of the increasing population in mixed urban-rural areas when moving from Small Areas (79,100 people) to Super Output Areas (160,900 people) and former Wards (206,300 people).

Lookup to drive times

21. A new addition to the settlement classification was the introduction of drive times¹. In the report, shading was applied to [maps](#) to demonstrate the area within a certain drive time of town centres within settlements with a population of 10,000 people or more. These maps have been overlaid with those of Small Areas and Super Output Areas, so that the banded drive time of the majority of the population in that area to a town centre can be presented. This information has been added to the lookup tables, firstly, by showing whether there is any town centre within 20 or 30 minutes drive time, and secondly, by showing banded drive times to 25 town centres (0-20 minutes, 20-30 minutes and more than 30 minutes). For Belfast and Derry City, there are five drive time bands with the 'more than 30 minutes' group split into 30-45 minutes, 45-60 minutes and more than 60 minutes.

¹ See <http://www.nisra.gov.uk/archive/geography/review-of-the-statistical-classification-and-delineation-of-settlements-march-2015.pdf#page=12>

22. [Table 1](#) in the settlement classification report shows that approximately 93 per cent of Northern Ireland's population live within 30 minutes drive-time of a town centre within a settlement containing a population of 10,000 people or more². The equivalent figure for 20 minute drive time is 80 per cent of the population. These figures were derived based on the geographical centroid of Small Areas, i.e. the drive time from the centre of the Small Area to the nearest town centre of a settlement containing a population of 10,000 people or more. These percentages are smaller than those based on exact boundaries and exact location of the population (95.5% and 83.7% respectively), however, this information was not available at the time of the settlement classification report.
23. The new lookup tables are created based on the majority of the population approach outlined earlier. For the bulk of Small Areas, the banded drive times were no different from those based on the geographical centroid. However, there were more Small Areas for which the drive times were overestimated than underestimated. The derived percentages of 95.6 per cent and 83.8 per cent respectively from the lookup table are very close to that based on exact boundaries and locations.
24. The settlement classification report ([Table 2](#)) provides an indication for each settlement with a population less than 10,000 people whether its population-weighted centroid lies within 20 or 30 minutes drive time of the town centre of a larger settlement. This does not necessarily mean that this applies to the whole population of that settlement. Indeed, several of these settlements are approximated by Small Areas which have a mixture of banded drive times.

Issues

25. The guidance provided in this paper is based on **information available at the time** of writing (i.e. May 2016). Going forward, clearly there will be changes in the population, the settlement development limits and the road network that will invariably give rise to different results in both the settlement classification and the lookup tables, including the Central Postcode Directory. Users should be aware of this possibility.
26. The advice and guidance **relates to count data only**. The guidance does not apply to percentages, rates or other derived statistics. Users wishing to create such statistics will need to apply this guidance to the baseline count data before creating the derived statistics.

² Note that there were no identified town centres in Metropolitan Castlereagh and Metropolitan Newtownabbey. There is a town centre in Lisburn, but not in the distinctly different 'Metropolitan Lisburn'. Craigavon Urban Area has three centres – Craigavon, Lurgan and Portadown.

27. It is recommended that any statistics produced for the new settlements should be accompanied by **metadata** outlining the method of spatial classification. The lookup table to approximate settlement statistics from Small Area statistics also provides an indication of accuracy using the 2011 Census counts of usual residents and occupied households. However, these indications of accuracy may not be applicable, particularly when reporting on sub-sections of the population or locally concentrated groups (for example, the elderly, ethnic minority groups, etc).

28. Data producers should always assess the risk of disclosing information on individuals, households, or businesses. The protection of the privacy of individual information whilst balancing the need for effective use of the statistics is embedded in the Code of Practice for Official Statistics³. If official statistics are produced from original micro data for multiple geographies, this can lead to **disclosure risks**. Producers should be particularly aware of this in any consideration of requests for statistics across multiple geographies, where information on cross-sections of geographies can be deducted through differencing. A producer may choose a less accurate method for one or more of the geographical areas to reduce this risk, if this would be more preferable than not releasing the statistics at all. The main statistical disclosure control techniques are described in

http://www.ninis2.nisra.gov.uk/public/documents/NINIS_Disclosure_Policy.doc

29. **Further guidance** is available from the NISRA Geography website⁴ or by e-mail from ninis.nisra@finance-ni.gov.uk.

Northern Ireland Statistics and Research Agency

May 2016

³ See <http://www.statisticsauthority.gov.uk/assessment/code-of-practice/code-of-practice-for-official-statistics.pdf>

⁴ See <http://www.nisra.gov.uk/geography/home.htm>

Annex A: 2011 Census population by Settlement Band – exact boundaries vs. approximations

	SETTLEMENT BANDS							
	A	B	C	D	E	F	G	H
Exact boundary⁵	280,211	83,125	522,217	128,937	117,434	77,423	103,500	498,016
Grid Squares⁶	280,138	83,163	522,382	128,861	117,438	77,534	103,539	498,016
<i>Difference with exact</i>	0.0%	0.0%	0.0%	-0.1%	0.0%	0.1%	0.0%	0.0%
Central Postcode Directory⁷	280,346	83,264	522,306	128,452	117,204	77,700	103,546	498,045
<i>Difference with exact</i>	0.0%	0.2%	0.0%	-0.4%	-0.2%	0.4%	0.0%	0.0%
Small Areas	280,962	83,403	525,649	130,130	119,726	80,976	109,555	480,821
<i>Difference with exact</i>	0.3%	0.3%	0.7%	0.9%	2.0%	4.6%	5.9%	-3.5%
URBAN SETTLEMENTS					RURAL SETTLEMENTS AND OPEN COUNTRYSIDE			

⁵ See <http://www.ninis2.nisra.gov.uk/Download/Census%202011/Headcount%20and%20Household%20Estimates%20for%20Settlements.ods>

⁶ Available at [http://www.ninis2.nisra.gov.uk/Download/Census%202011/KS102NI%20\(Se\).ods](http://www.ninis2.nisra.gov.uk/Download/Census%202011/KS102NI%20(Se).ods)

⁷ Based on the Central Postcode Directory of November 2015 (see <http://www.nisra.gov.uk/geography/postcode.htm>) and the 2011 Census headcount and household estimate by postcode (see <http://www.nisra.gov.uk/census/2011/results/postcode.html>).

**Annex B: 2011 Census population and household estimates for urban, rural and mixed urban-rural areas
– exact boundaries vs. approximations**

	TOTAL POPULATION			TOTAL HOUSEHOLDS		
	URBAN AREAS	MIXED AREAS	RURAL AREAS	URBAN AREAS	MIXED AREAS	RURAL AREAS
Exact boundary⁸	1,131,924	-	678,939	462,810	-	240,465
Small Area to urban-rural classification	1,087,724	79,052	644,087	446,481	28,845	227,949
<i>Difference with exact</i>	-44,200	+79,052	-34,852	-16,329	+28,845	-12,516
Super Output Area to urban-rural classification	1,047,088	160,857	602,918	430,657	59,531	213,087
<i>Difference with exact</i>	-84,836	+160,857	-76,021	-32,153	+59,531	-27,378
Ward1992 to urban-rural classification	1,020,336	206,273	584,254	420,423	76,253	206,599
<i>Difference with exact</i>	-111,588	+206,273	-94,685	-42,387	+76,253	-33,866

⁸ See <http://www.ninis2.nisra.gov.uk/Download/Census%202011/Headcount%20and%20Household%20Estimates%20for%20Settlements.ods>