Scotland's Census
Shaping our future
A' dealbhadh ar n-àm ri teachd

Scotland’s Census 2022

Cell Key Perturbation

October 2020

National Records of Scotland

Contents

1. Plain English Abstract

Some people in Scotland may have characteristics which make them rare or unique within their local area. In such cases, these people might be easy to identify in tables published from Scotland's Census 2022. National Records of Scotland (NRS) will therefore take steps to protect the privacy of people and households in Scotland and ensure that they cannot be identified in any published outputs (tables, reports etc.) from the census.

One of the methods used will be adding to or taking away a small number from some of the cells in a table, known as Cell Key Perturbation. This will mean that when users compare tables they can't be sure whether any differences are between tables are real people, or are caused by this Cell Key Perturbation.

2. Abstract

Statistical Disclosure Control (SDC) describes methods applied to census data and outputs to protect the privacy of personal information. It can include making small changes to data, controlling access to data, and controlling the detail that is available to census data users.

In 2022 we are adding a flexible table builder to our range of output tools, to allow users to create tables that meet their needs exactly. The inclusion of this new output method requires and additional layer of SDC called 'Cell Key Perturbation'. This is to protect against differencing and allow users to create their own tables. This method is already used by other statistics organisations such as the Australian Bureau of Statistics, and the Department for Work and Pensions.

Cell Key Perturbation adds a small amount of noise to some cells in a table, meaning that users cannot be sure whether differences between tables represent a real person, or are caused by the perturbation. Cell Key Perturbation is consistent and repeatable, so the same cells are always perturbed in the same way.

In addition to this, ONS are proposing to also perturb zeros in order to add uncertainty as to whether a 1 in a table represents a 'real' person.

Perturbation of zeros is repeatable, but not consistent across tables in the same way that Cell Key Perturbation is. This could make it more difficult for users to understand the data, and raise concerns over perceived data quality.

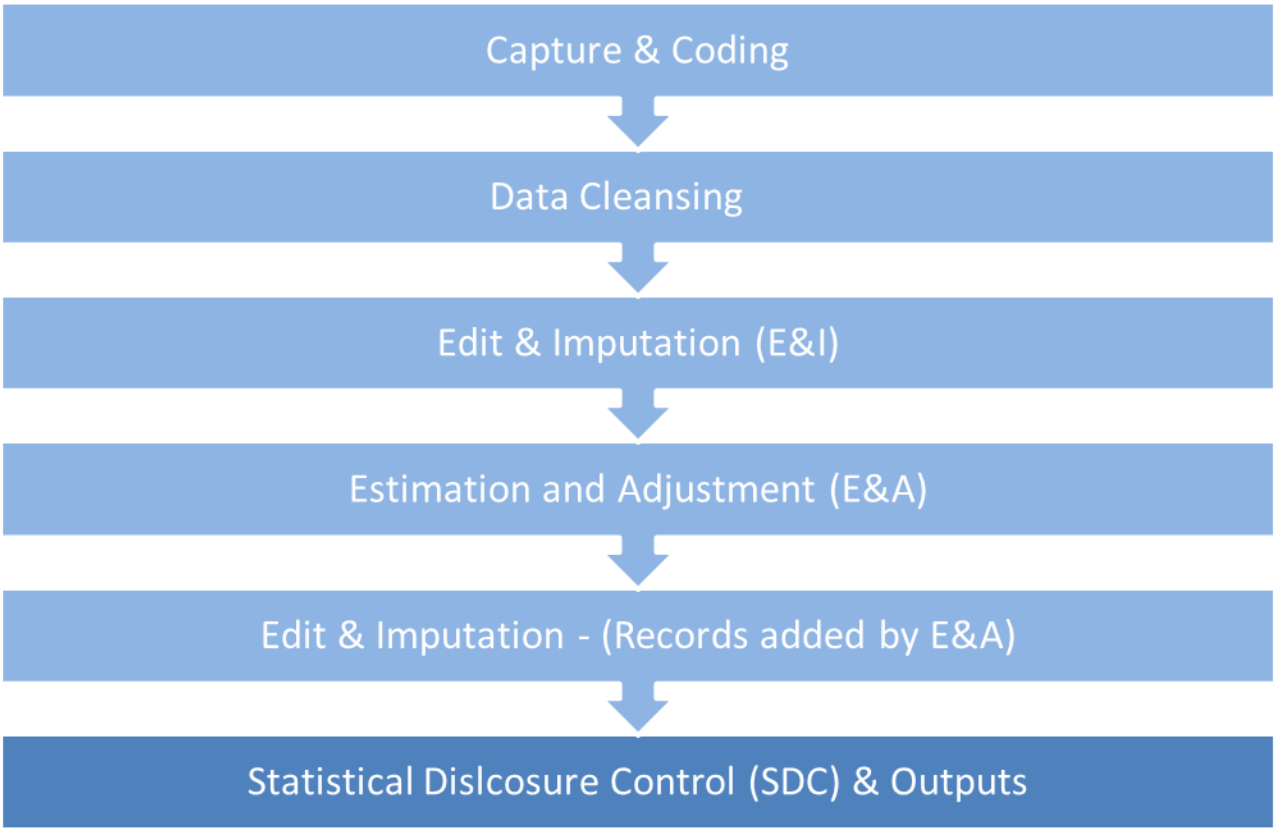
We are recommending that we add cell key perturbation but not perturbation of zeros to the SDC methodology for 2022.

3. Introduction and Background

NRS has a legal obligation to ensure that the privacy of individuals and households in Scotland is protected in all outputs from Scotland's Census 2022. For more information on confidentiality protection in the context of the census and relevant legislation, please see the Confidentiality page on the Scotland's Census website.

In order to prevent the release of confidential information, NRS will utilise a number of Statistical Disclosure Control (SDC) techniques. SDC refers to the range of methods which can be used to prevent the release of confidential information about an individual or household in census outputs.

The application of statistical disclosure control techniques is the final step in the census statistical methodology prior to the creation and dissemination of outputs. For more information on the full statistical methodology for Scotland's Census 2022, please refer to the Statistical Methodology page on the Scotland's Census website.



Cell Key Perturbation is a post-tabular method of SDC, i.e. it is applied to tables, rather than to the microdata.

In 2011 targeted record swapping[1] was applied to the census microdata. Subsequently, each table was checked for disclosure. Tables were checked to ensure they had 'sufficient uncertainty', and thus were safe to release. Part of this

[1] Link to Record Swapping Paper

checking included differencing tables, to make sure that no extra information could be gained by comparing two published tables.

Age by Sex	0-4	5-9	10-14
Male	9	15	21
Female	9	13	20

Age by Sex	0-5	6-10	11-15
Male	10	17	18
Female	11	15	17

Figure 1: Original Tables

Age by Sex	5	10	15
Male	1	3	0
Female	2	0	1

Figure 2: Data Calculated by Differencing

Figure 1 shows two tables are considered safe to release when considered alone, but when compared (Figure 2), they allow the user to learn extra information that would not be considered safe to release. Therefore, all tables had to be checked against all other published tables, to ensure that users could not combine them to learn additional information that would not be published.

This process took considerable time, and led to delays in the release of outputs. In addition, as this limits which tables can be produced, users may have had to accept a table that did not meet their needs exactly.

To address these issues, Cell Key Perturbation is being added as an extra layer of SDC, allowing users to create their own tables, while protecting against differencing.

4. Cell Key Perturbation Methodology

Each record in the microdata is assigned a random number record key. These keys are then added together to get the cell key for each cell in a table.

The perturbation table (p-table) determines how much each cell in a table is perturbed by. Most cells in the p-table are zero, representing no change to the cell, but a small number of cells will contain -2, -1, 1, or 2. The number of cells in a p-table with non-zero values is determined by the perturbation rate (for example a perturbation rate of 10 would mean that 10% of cells in the p-table were non-zero). The total sum of all the cells in the p-table is zero, so that we are not introducing bias to the outputs.

1) Assign each record a random number (record-key)

Record	Rkey
r1	54
r2	104
r3	93
…	…
rN	26

2) Create a frequency table. For each cell, sum record keys and take the modulo to get the cell key

Age by Sex	Male	Female
0-15	.	.
16-24	.	4
25-34	.	.
…		

Record	Rkey
r2	104
r4	61
r56	7
r72	90
Sum Rkey = 262	
Cell key = 262 mod 200 = 62	

3) Use perturbation table to get perturbation value from cell value and cell key

Ptable		Cell key (1-200)						
		1	2	3	…	62	…	200
Cell Value	1		+1					
	2	-1						
	3							-1
	4	+1				+1		
	5			-1				
	…							

4) Apply the chosen perturbation to the cell

Age by Sex	Male	Female
0-15	.	.
16-24	.	5
25-34	.	.
…		

Figure 3: Cell Key Perturbation Method

The cell key is used with the p-table to add a small amount of noise to some of the cells in the table. As the cell key is created using the records in each cell, this is replicable and the table will always be perturbed in the same way. The same records appearing together in different tables will have the same cell key value, so will always be perturbed in the same way.

Figure 3 shows how the Cell Key Perturbation method is applied to tables.

The small changes to some cells means that users cannot be sure whether the difference between two similar tables is a real difference, or due to the cell perturbation. Some of the differences obtained may be illogical, but this will let users know that perturbation has been applied and that counts obtained by differencing are unreliable. Figure 4 and 5 show and example of this, where by differencing two tables, a user would get a difference of -1.

Age by Sex	0-4	5-9	10-14
Male	8	15	21
Female	9	13	22

Age by Sex	0-5	6-10	11-15
Male	10	17	18
Female	11	15	17

Figure 4: Original Tables

Age by Sex	5	10	15
Male	2	3	0
Female	2	0	-1

Figure 5: Data Calculated by Differencing with Perturbation

5. Testing the Method

We carried out internal testing, to assess the method and whether it made the data more secure, i.e. harder for users to learn extra information that had not been published. We first carried out some checks within the team to identify an acceptable level of perturbation, that is a level that protected the data, without substantially damaging the utility. We then applied this level of perturbation to a range of tables to carry out 'unpickability' testing.

5.1 Determining the level of perturbation

In order to assess how easy it would be to unpick the perturbation, we looked at the different types of perturbation that could occur in a table:

- one – only one cell in a table has been perturbed, making it easy to unpick
- matching totals - the perturbation matches the difference in the column and row totals of the table
- multiple perturbations – multiple perturbations of the same size has taken place, making it uncertain which cells were perturbed
- multiple same inconsistencies – multiple inconsistencies of the same size, in the table and on column and row totals
- multiple perturbations on total – multiple perturbations on the same row or column

These perturbations were classed as 'bad' or 'good', with one, and matching totals considered 'bad' perturbation (easy to unpick), and the others considered 'good' perturbation (harder to unpick).

A selection of tables were perturbed, and the different types of perturbation measured. This allowed us to analyse the point at which there was more 'good' perturbation than 'bad' perturbation, and use this to decide the perturbation rate to test.

5.2 Testing

Tables of data on religion, ethnic group and country of birth were produced by age and sex at datazone[2] and output area level. Internal testing was carried out by asking members of staff to act as users and try to identify which cells in these tables had been perturbed. The table below shows how often these perturbations were correctly identified.

Sample Table Reference	Total number of cells perturbed	Perturbations correctly identified	Perturbations missed	Incorrectly identified perturbations

[2] See details of NRS geographies here

A1	18	11	7	3
A2	18	13	5	5
B	30	19	11	6
C1	46	29	17	14
C2	46	22	24	11
C3	46	13	33	31
D	20	15	5	5
E	36	9	27	0

Figure 5: Table of Unpickability Testing Results

From the testing carried out there were no cases where users could correctly identify all cells that had been perturbed. This small scale testing demonstrates that cell perturbation is difficult to unpick so would be an appropriate additional method of SDC to use to protect against differencing.

Using this method, deriving any additional information through differencing would require a user to unpick all of the perturbation in multiple tables. There is a very low probability of effectively unpicking the data in this way. In addition, users have no way of knowing whether the perturbation was unpicked fully and correctly.

Record swapping and imputation will be carried out on the data prior to cell perturbation, so even if users could unpick the perturbation the data will still be protected. Additionally, users would have to agree to the terms of access when using the flexible table builder, which would include an agreement not to attempt to unpick perturbation.

5.3 Limitations/Issues

Although the Cell Key Perturbation process is consistent and repeatable, it can introduce inconsistencies in the data - commonly the cells in a row or column will not add up to the total displayed. This can create some confusion for uses as they are not sure which to trust as the ‘real’ number.

Another issue is that summing together areas to a higher geography, could give different results due to the perturbation applied. For example, adding together all of the datazones in a council area, would not give the same numbers as that table produced at council area.

These issues will require clear messaging from NRS to explain the minor inconsistencies in the data, and the best way to use the table builder to get the data needed.

6. Perturbation of Zeros

6.1 Method

Regular Cell Key Perturbation uses record keys and the value of the cell to determine which cells to perturb. As zero cells do not have any records associated with them, a different method needs to be used to carry out the perturbation. The Office for National Statistics (ONS) is proposing to add the perturbation of zeros to its SDC methodology for the next census.

To perturb zeros, every category of a variable is assigned a category key, that will be used to determine which zero cells to perturb (Figure 5).

Variable	Entry	Category Key
Sex	1	1980
Sex	2	7078
Age	0-5	3839
Age	6-10	9954
Age	11-15	6492
Age	16-20	8291

Figure 6: Example of category keys

For each zero cell in a table the corresponding category keys are added together (e.g. sex + age). Similarly to the record keys, these category keys would be generated once and remain the same, meaning that the same cell in the same table will always be perturbed in the same way.

In order to prevent impossible combinations from being generated, only cells with a '1' at Local Authority level will be perturbed. This means that perturbation of zeros will only occur if there is a unique record at Local Authority level. For example, in Figure 7 below, there is only one record aged 10 to 14 with the religion Buddhist. If this table was created at a lower geography within this Local Authority (e.g. Datazone), then this cell could be perturbed, but zeros in other cells would not.

Age by Religion	Buddhist	Hindu
0 to 4	2	10
5 to 9	0	6
10 to 14	1	6

Figure 7: Local Authority Table

All of the cells with a category key above a certain value are then perturbed from 0 to 1. This value is determined by the perturbation rate. A corresponding number of cells with a value of 1 will be perturbed down to 0, to prevent introducing bias.

6.2 Unpickability Testing

In order to evaluate whether perturbing zeros made it harder to unpick the Cell Key Perturbation, a selection of commonly used tables were built using regular Cell Key Perturbation, and the same tables using cell key perturbation, including perturbation of zeros.

Perturbations were correctly identified between 50 and 70 per cent of the time in tables without perturbation of zeros, and between 50 and 73 per cent in tables with zero perturbation applied. Overall the level of unpickability was broadly similar between the same tables with and without zero perturbation applied. The notable exception to this was tables that included ethnic group, where the unpickability level rose by up to 10 percentage points. This is likely driven by the fact that there is less diversity in ethnic group data than other variables, and, therefore, more 1s at Local Authority level, and a greater number of cells with the potential to be perturbed. For comparison, other tables had fewer than 15 1s at the Local Authority level, whereas ethnic group data had 41.

The purpose of perturbing zeros is to add uncertainty as to whether a 1 appearing in a table represents a 'real' person. Without perturbation of zeros, if a user sees a 1 in a table, they can be sure that there is at least one person with that combination of characteristics in a given area. If zeros are perturbed a user cannot be certain whether that combination is present in a given area.

6.3 Issues

One of the main concerns over perturbing zeros is that while it is repeatable, it is not consistent across tables.

	0-5	6-10	11-15	16-20
Male	0	2	3	2
Female	1	2	4	3

	0-10	11-20
Male	2	5
Female	3	7

Figure 8: Perturbation across tables

	0-5	6-10	11-15	16-20
Male	0	1	3	2
Female	1	2	4	3

	0-10	11-20
Male	0	5
Female	3	7

Figure 9: Perturbation of zeros across tables

In Figure 8 above, the highlighted cell has been perturbed from 1 to 2. In all tables where this record appears, it will always be perturbed to 2. For example, changing from 5 year age bands to 10 year age bands, or adding an additional variable to the table will always result in the same perturbation being applied.

The same does not hold true for perturbation of zeros. As the perturbation of zeros is based on the variables and the variable categories, a slightly different breakdown, or an additional variable will change what cells the perturbation is applied to. This is highlighted in in Figure 9, where changing from 5 year age bands to 10 year age bands does not carry across the 1. This may make it more obvious to users where perturbation has been applied.

We also need to consider how users would perceive this change. Whether it would appear that we are adding an extra barrier to understanding, with no clear benefit to the user, and how would this affect perceived data quality.

Another concern is whether we would be able to give users details of the perturbation of zeros methodology. While we do not share specific details with users, such as swapping variables, or perturbation rate, we do share details of the methodology and explanations of how they work.

While perturbation of zeros may work well for ONS, the demography of Scotland differs from that of England and Wales, meaning the perturbation of zeros may have a more damaging effect on Scottish data. In general Scotland is less diverse than England and Wales, for example in the 2011 census 92 per cent of the Scottish population identified their ethnicity as White British[3], compared with 84 per cent of the English and Welsh population. This, combined with the much smaller population of Scotland when compared to England and Wales, and difference in geographies – see Appendix A between the two countries, may mean that areas in Scotland are disproportionally affected by perturbation of zeros, and it would have a much more damaging effect on Scottish data.

7. Conclusion

Overall, we are proposing to add Cell Key Perturbation to our SDC methodology for Scotland's Census 2022, but not perturbation of zeros.

The addition of the Cell key Perturbation methodology in 2022 will protect against differencing, while allowing users to create their own tables. Record swapping targets unique individuals or households so this remains the main method of disclosure control, with Cell key Perturbation being an addition to protect against differencing.

At this time we feel that there are still concerns around the perturbation of zeros, particularly in terms of how it may affect data in Scotland. Therefore, we are not including the perturbation of zeros in our SDC methodology for the 2022 census.

[3] In Scotland this includes White Scottish, and White Other British.

8. Appendix A – Small area statistical geographies in the UK by country

2011 Census population and household counts by statistical geography

Country and small area geography	Number	Population			Households		
		Minimum	Average	Maximum	Minimum	Average	Maximum
England and Wales: Output Areas[1]	181,408	91	309	4,140	23	129	817
England and Wales: Lower Layer Super Output Areas[2]	34,753	983	1,614	8,300	304	672	1,405
England and Wales: Middle Layer Super Output Areas[3,4]	7,201	5,003	7,787	16,342	2,003	3,245	6,100
Scotland: Output Areas	46,351	50	114	2,081	20	51	126
Scotland: Data Zones	6,976	147	759	2,901	36	340	817
Northern Ireland: Small Areas	4,537	98	399	3,072	59	155	988
Northern Ireland: Super Output Areas	890	364	2,035	4,574	200	790	1,698

[1] Target was for a minimum of 100 and a maximum of 625 people, and a minimum of 40 and a maximum of 250 households.
[2] Target was for a minimum of 1,000 and a maximum of 3,000 people, and a minimum of 400 and a maximum of 1,200 households.
[3] Target was for a minimum of 5,000 and a maximum of 15,000 people, and a minimum of 2,000 and a maximum of 6,000 households.
[4] Minimum figures exclude the MSOA covering the whole of the Isles of Scilly, which had much lower population and household counts.

Please note - the need to include communal establishment residents within the smallest output geographies (Output Areas/Small Areas) may result in maximum population thresholds not being achievable for some areas. In such cases the upper target population threshold will intentionally be breached.

For each country, the small area geographies are hierarchical:
For England and Wales, OAs nest within LSOAs, LSOAs nest within MSOAs, and MSOAs nest within local authorities.
For Scotland, OAs nest within Data Zones, and Data Zones nest within council areas.
For Northern Ireland, Small Areas nest within SOAs, though neither Small Areas nor SOAs wholly nest within the current district council areas, but did for the district council areas as constituted at the time of the 2011 Census.