Statistical Disclosure Control and Outputs for the 2021 Census

National Records of Scotland

Scotland's Census
Shaping our future

Overview

- **Statistical Disclosure Control**
- Webpage Mockups
- SuperWEB2
- Q&A

What is statistical disclosure control?

Statistical Disclosure Control (SDC) is required to:

Prevent the release of any information about an individual, household or enterprise that involves or could lead to:

- their identification, or
- the disclosure of confidential information about them

What is statistical disclosure control?

SDC involves

- either:
  - Introducing sufficient ambiguity/uncertainty into, or reducing the level of detail of published statistics so that the risk of disclosing **confidential information** is reduced to an **acceptable level**

- and / or:
  - Controlling access to data

National Records of Scotland

Scotland's Census
Shaping our future

Why do we need SDC?

- Legal
  - Census Act 1920
  - Census (Confidentiality) Act 1991
  - Data Protection Act 1998
  - The Census (Scotland) Regulations 2010

- Ethical
- Practical

What we did for the 2011 Census

- Targeted record swapping
  - Targeted to "risky" records

- Table redesign
  - Criteria of % 1s and attribute disclosures that are "real"

Every table had to be checked for disclosure
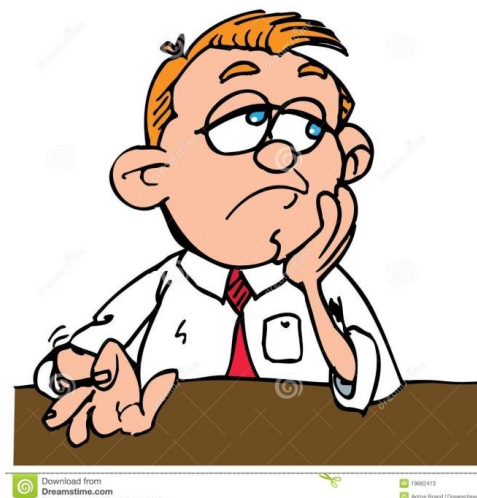
Timing was affected

Targeted Record Swapping

- Select a **sample** of records/households to be swapped

- Using a set of variables, find a **match** for each sample record/household

- **Swap** the geographic variables of sample record/household with that of matched record/household

Characteristics:
Age: 22
Sex: Male
Marital Status: Single
No of Cars: 3
Region: Area A

A

B

Characteristics:
Age: 22
Sex: Male
Marital Status: Single
No of Cars: 1
Region: Area B

National Records of Scotland

Scotland's Census
Shaping our future
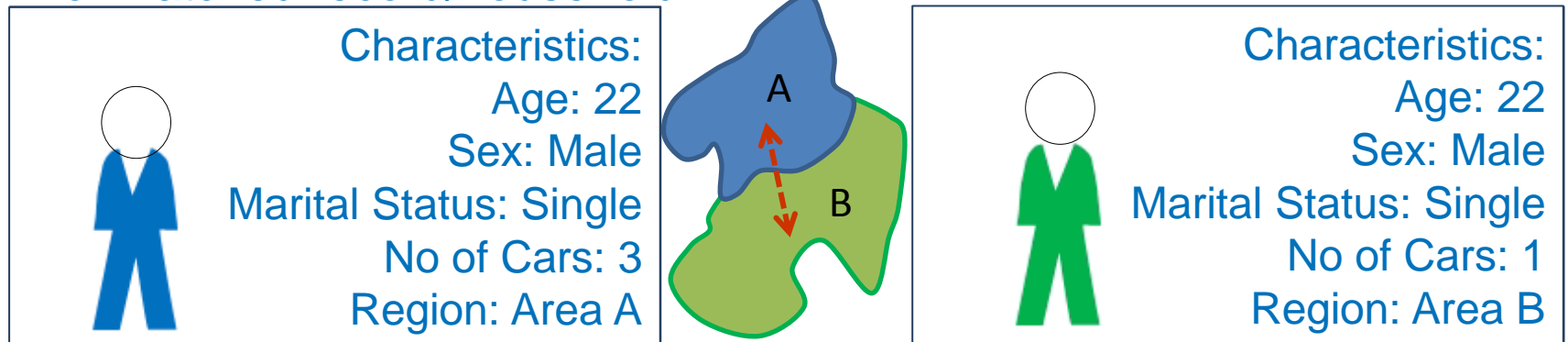
What did we learn from the 2011 Census?

- The process of SDC for the 2011 Census was not as efficient as it could have been:
  - Each table had to be manually checked for disclosure risks, taking up a lot of time an resources
  - Manual manipulation of all tables was required to ensure no disclosure issues occurred
  - At times, this caused lengthy delays in the release of publications and the release of data for customers

This resulted in less than ideal outputs

Sometimes I wanted this table…

They gave me this table

2021 Census

- For the 2021 Census a we are planning to build on what was done in 2011

- Planning to use targeted record swapping

- We are planning on using a method of Cell Key Perturbation

- This should:
  - Reduce the time from gathering the data to when we publish data
  - Reduce the need of manually checking tables which would speed up the publication process

- Aim to release some tables at higher geographies unperturbed

National Records of Scotland

Scotland's Census
Shaping our future

Cell Key Perturbation

1 Assign each record a random number

Record	Rkey
$r_1 \rightarrow$	54
$r_2 \rightarrow$	4
$r_3 \rightarrow$	93
…	
$r_N \rightarrow$	26

2 For each cell, sum rkey and apply a function to get a cell key

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

Record	Rkey
$r_2 \rightarrow$	4
$r_4 \rightarrow$	61
$r_{56} \rightarrow$	7
$r_{72} \rightarrow$	90
Sum =	162

e.g. take last two digits → **Ckey = 62**

3 Use a look up table to get perturbation value

Cell Key →

Cell Value ↓

	1	2	3	…	61	62	63	…	99
1		+1							
2			+1				-1		
3									+1
4	-1					**+1**			
5			-1		-1				
…									

4 Apply pvalue to cell

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

Notes for Cell Key Method

- Adapted from "Australian Bureau of Statistics (ABS) method"

- Primarily a protection against 'differencing'

- Looking at a light touch (record swapping still the primary approach)

- Considering the need to retain 1s and 2s in outputs, by perturbing zeros (KEY PART OF THE PROTECTION)

- Introduces another layer of uncertainty for intruder

- Consistency in same cell across tables

- Some inconsistencies in breakdowns

SDC Questions

- What do you like/dislike about our proposed 2021 SDC methods?

- Do you foresee any problems or issues with the proposed 2021 SDC methods?

- In your opinion is the trade-off of more timely outputs and increased flexibility worth the addition of cell key perturbation.

- Will unperturbed tables be beneficial? If so, which higher level geographies are required.

Overview

- Statistical Disclosure Control
- **Webpage Mockups**
- SuperWEB2
- Q&A

New 2021 Census webpages

- We are aiming to produce an updated website/webpages to coincide with 2021 Census data.
    - Incorporate a new user interface
    - Overhaul the website meaning a more streamlined user experience and modern look
    - New functions and features may be added to improve the workflow and overall experience
        - Integrated Metadata
        - Utilisation of Infographics and pre-defined tables
        - Mobile device functionality
    - More functionality in creating tables by utilising a flexible table generator
        - Encouraged to create a login to get access to extra options and functions

National Records of Scotland

Scotland's Census
Shaping our future

2021 Webpages – Home Page

2021 Webpages – Visualisation

2021 Webpages – Mobile View
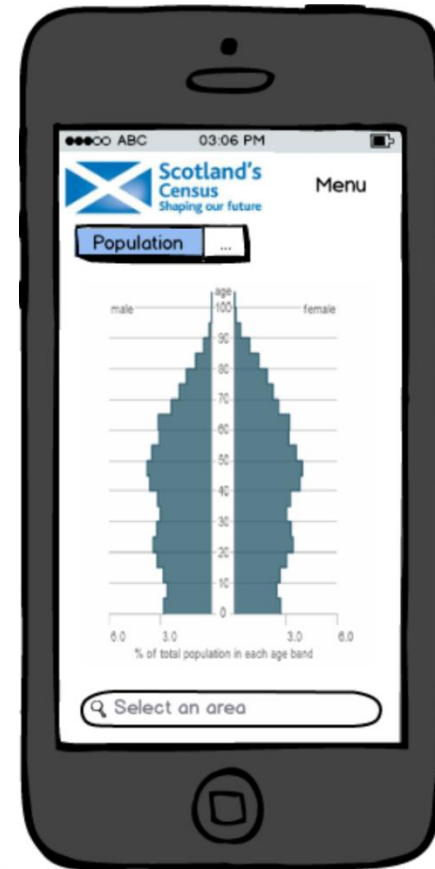
2021 Webpages – Tablet View
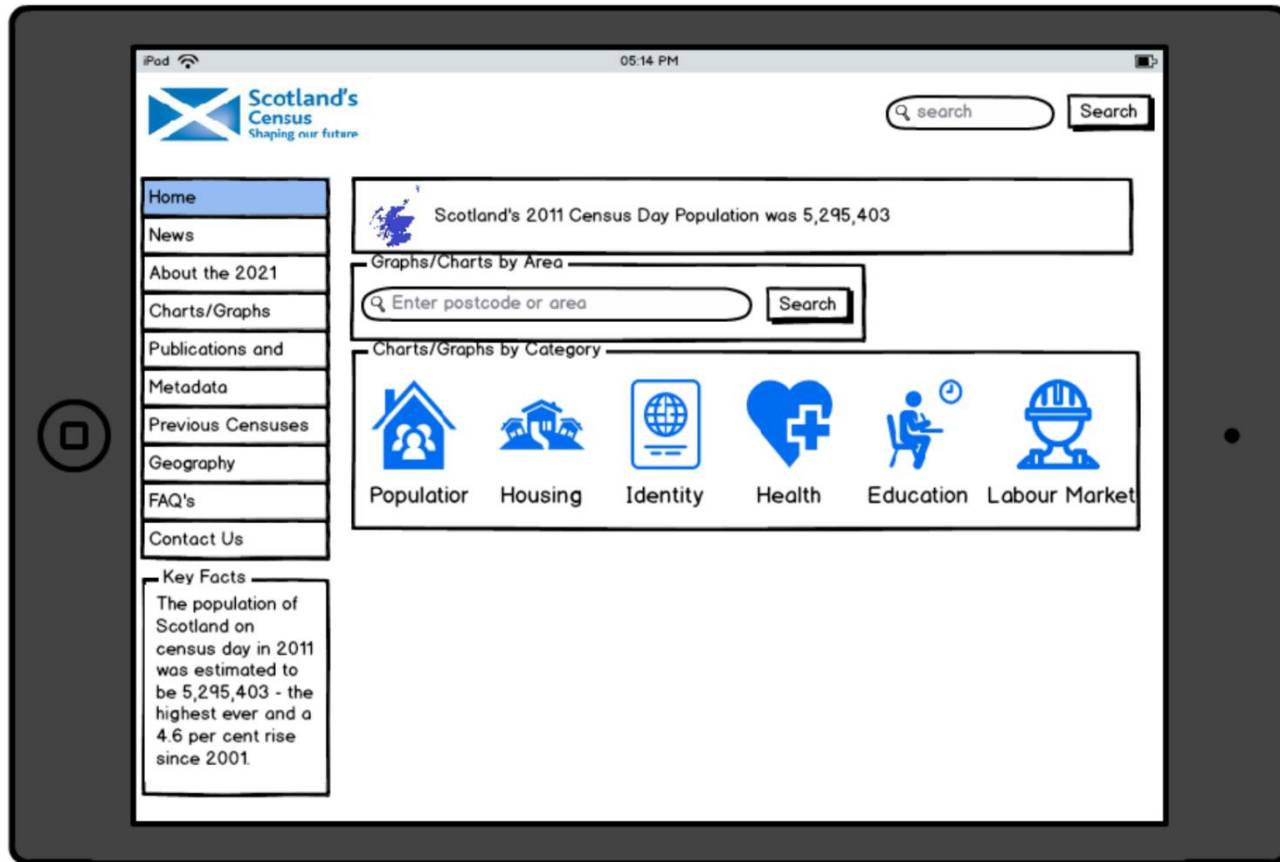
2021 Webpages– General Health Table

2021 Webpages– Table Generator

2021 Webpages– Differing Geography View

2021 Webpages– SDC Pop-up

2021 Webpages– SDC Pop-up

Overview

- Statistical Disclosure Control
- Webpage Mockups
- **SuperWEB2**
- Q&A

SuperWEB2

While using the SuperWEB2 package think over these questions:

- How easy is it to use?

- Are all desired functions available?

- What would you like to see added?

- Were there any issues with the software?

- Would you be happy to use this type of software to create your own tables?

Next Steps

- After the event we will email you with a
  - Link to the slides from this event on our website.
  - Link to a questionnaire to obtain more detailed feedback on your views of our SDC and output plans.
- We will then analyse all the feedback we receive from this event and from the questionnaires
- In August/September we plan to:
  - Publish a summary of all the responses we have received
  - Release synthetic data (based on the structure of the 2011 Census) with our existing flexible table generator.

National Records of Scotland

Scotland's Census
Shaping our future

Finally

- Thanks for all your contributions today

- Please complete our short feedback form before leaving

- If you have any further questions please speak to us or email scotlandscensus@nrscotland.gov.uk