

# Postcode Anonymity Analysis

Adam Hardy

May 2021

## 1 Introduction

In the US, 5-digit Zip codes are usually rounded to 3-digits when anonymising healthdata, so knowledge of the Zip code doesn't allow small groups to be identified. Even then, there are some 3-digit codes that have fewer than 20,000 residents, and the advice is to lump these together under a new code (000). Looking forward to how GDPR may affect data handling in the UK, is it possible to use a similar approach here?

Below, a list of UK postcodes and population data<sup>1</sup> was used for an anonymity analysis on UK postcodes at different geographic levels and population groups.

UK postcodes can be in one of six formats and broken down into 7 components that describe geographic areas (Table 1). We will use the area, district, sub-district and sector. Where a postcode does not have a subdistrict, we will substitute the district in its place.

## 2 Anonymity Analysis

### 2.1 Postcode Component Selection

Given how UK postcodes are constructed, we can't simply truncate postcodes to 3 characters as is done for USA zip codes as this will result in postcode areas being combined together in ways which do not have real geographic meaning. For example, 'NE3 1ED' and 'NE35 2FG' are both valid postcodes. Truncating both of these to the first three characters would put these both in the 'NE3' group. They should instead be both in an 'NE' area group or separate 'NE3' and 'NE35' district groups (or sector or sub-district groups).

---

<sup>1</sup>[http://www.nomisweb.co.uk/output/census/2011/Postcode\\_Estimates\\_Table\\_1.csv](http://www.nomisweb.co.uk/output/census/2011/Postcode_Estimates_Table_1.csv)

Postcode Format	Outward Code	Inward Code	Area	District	Sub-District	Sector	Unit
AA9A 9AA	AA9A	9AA	AA	AA9	AA9A	AA9A 9	AA
A9A 9AA	A9A	9AA	A	A9	A9A	A9A 9	AA
A9 9AA	A9	9AA	A	A9	N/A	A9 9	AA
A99 9AA	A99	9AA	A	A99	N/A	A9 9	AA
AA9 9AA	AA9	9AA	AA	AA9	N/A	AA9 9	AA
AA99 9AA	AA99	9AA	AA	AA99	N/A	AA99 9	AA

Table 1: UK postcode formats and geographic components. 'A' represents an alphabetic character, '9' represent a numeric character. Obtained from <https://ideal-postcodes.co.uk/guides/uk-postcode-format>.

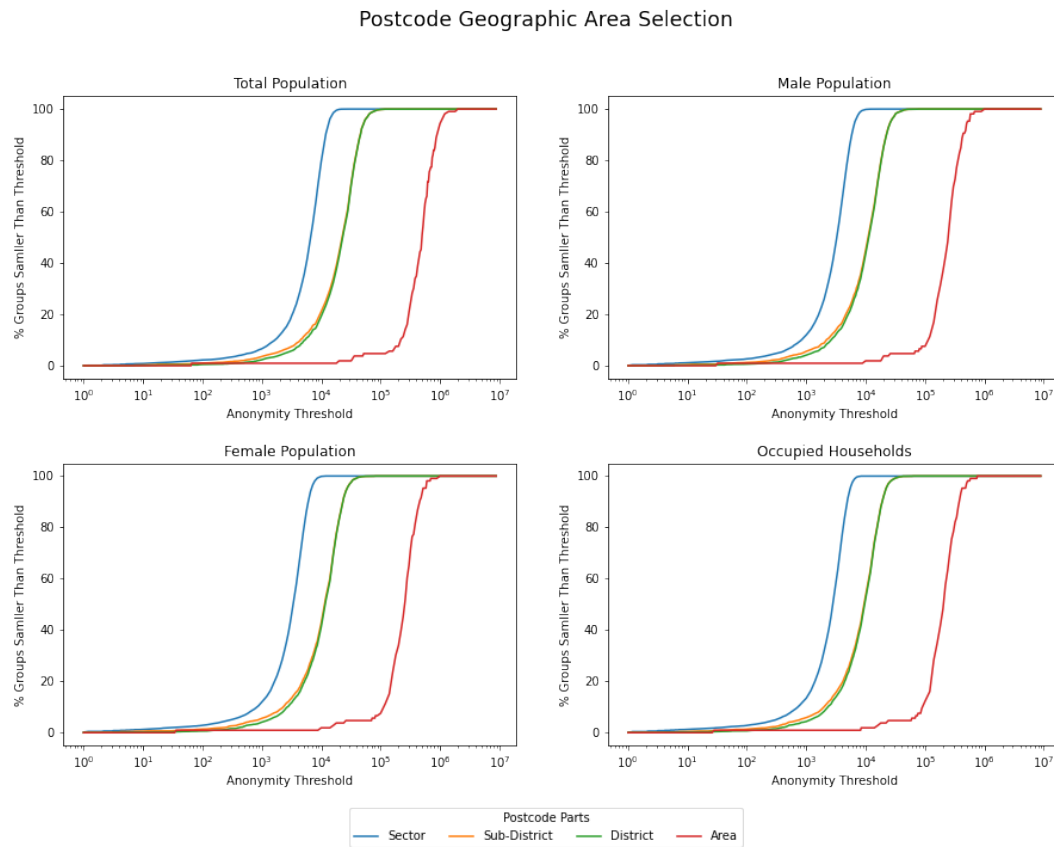


Figure 1: postcode selection

## 2.2 Population Segment Selection

## Population Selection

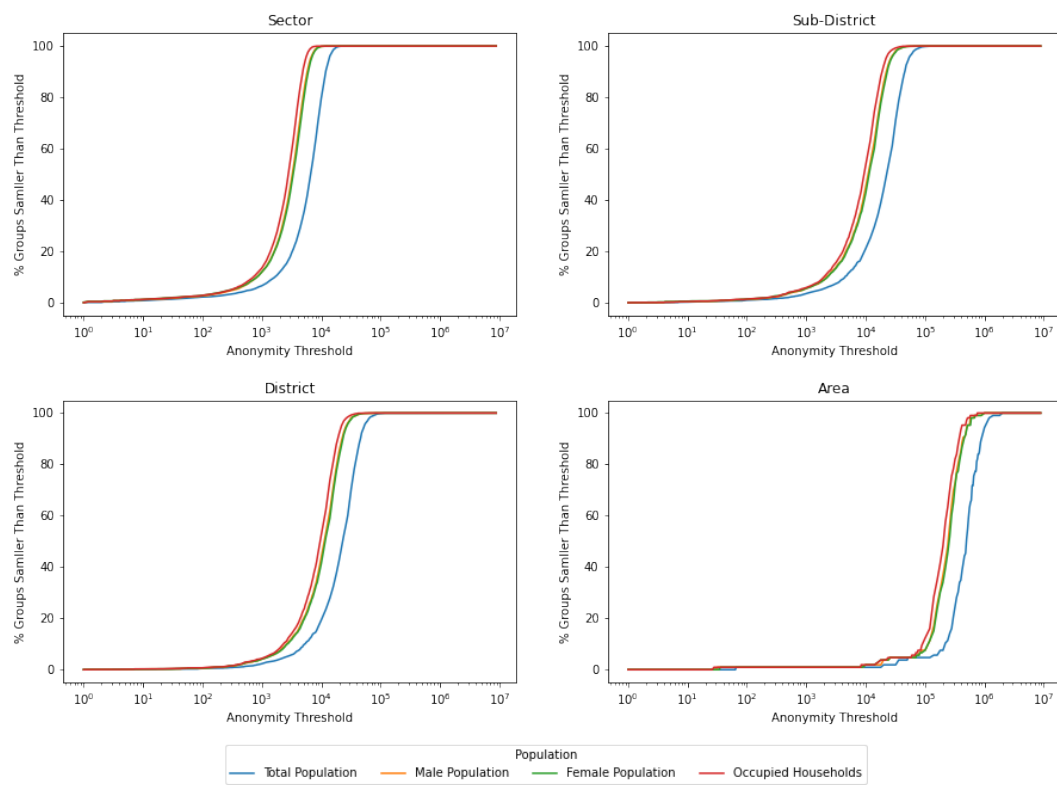


Figure 2: population selection