

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

telephone [REDACTED]  
email owen.boswarva@gmail.com

18 April 2022

The Information Commissioner's Office (ICO)

By email to casework@ico.org.uk

Department for Transport reference: F0020605

**Concern about accessing or re-using information from a public body –  
FOI and RoPSIR request to the Department for Transport for statistical information about  
the number of licensed vehicles in the UK**

Dear ICO,

I would like to report a concern about the Department for Transport's handling of an access to information and re-use of information request. A copy of my email correspondence with the DfT, including the internal review response, is enclosed in a separate PDF.

The information request

The DfT and the Driver and Vehicle Licensing Agency (DVLA) publish statistics for licensed vehicles by postcode district and body type (Table VEH0122) at:

<https://www.gov.uk/government/statistical-data-sets/all-vehicles-veh01>

That data is published for re-use under the terms of the Open Government Licence.

In December 2021 I requested from the DfT equivalent data (i.e. the number of licensed cars, motorcycles, and all other vehicles) by 2011 Census Output Area (OA) in England, Scotland and Wales, and Small Area (SA) in Northern Ireland. I requested this information without rounding or suppression of small numbers.

### The disclosed information

In January 2022, the DfT disclosed most of the requested information to me in an Excel workbook containing two tables of vehicle counts. I have enclosed a copy of the Excel workbook.

The first table contains 231,933 rows of data with vehicle counts for Output Areas (OAs) in the following format:

<u>OA11CD</u>	<u>Cars</u>	<u>Motorcycles</u>	<u>Others</u>
E00000024	80	5	29
E00042155	355	6	209
E00042425	119	c	6
E00042436	53	0	c
E00167301	152	8	16

The second table contains 42,619 rows of data Lower Layer Super Output Areas (LSOAs), in the same format. An LSOA is a larger Census output area that contains multiple OAs.

However, the DfT has not disclosed some of the information I requested. As explained in its response, the DfT has suppressed all values between 1 and 4:

“In instances where originally values would have been between 1 and 4, these data have been suppressed and the values in these cells replaced with 'c'. In addition to this, some values at LSOA-level have received secondary suppression to avoid the OA-level suppressed figures being disclosed using subtraction. This approach is consistent with all the statistical outputs that DfT produces from its extract of DVLA data.”

In response to my re-use request, the DfT said only:

“This dataset can be freely used and shared openly.”

### Summary of correspondence

In February 2022, I requested an internal review of the DfT’s response to my requests. I disputed the DfT’s application of the FOIA section 40(2) exemption to withhold values between 1 and 4. I also said that the DfT had not provided a clear statement of permission and/or conditions for my re-use of the disclosed information.

In March 2022, the DfT provided an internal review response that upheld its decision to apply the section 40(2) exemption.

With respect to my re-use request, the internal review response said the Open Government Licence v3 allows me to re-use the disclosed information, but also that the extract of data belongs to the DVLA and that the DfT statistics team responded to my information requests on behalf of the DVLA.

On 9 March 2022, I emailed the DVLA to seek its permission to re-use the information disclosed by the DfT. However, to date I have not received a response from the DVLA.

### Points of concern

The DfT maintains that disclosure of the suppressed values would contravene the Data Protection Act 2018:

“Due to the small number of vehicles registered in particular areas, release of information would be likely to lead to identification of individuals if pieced together with other information. The mosaic effect of combining information may potentially identify individuals who have a reasonable expectation that information about them is not placed into the public domain. As section 40 is an absolute exemption, there is no requirement to carry out a public interest test.”

In my view, the DfT is mistaken in its application of the section 40(2) exemption. The suppressed values are not personal data within the meaning of data protection law and the information cannot be used to identify any individual, either by itself or when combined with other available information.

Output Areas (OAs), the smallest area of geography in my information request, are a [Census geography](#) with a confidentiality threshold of 100 persons and 40 households. In 2011, the average OA had a population of more than 300 persons.

The nature of the vehicle licensing data is highly generalised, as the vehicle counts are by body type: car, motorcycle, or other vehicle.

Even if the vehicle licensing data showed that only a single vehicle was registered in a particular Output Area, that statistic could only be linked to an identifiable individual if it was already known from other information that the individual kept a vehicle within the Output Area. In that scenario, the statistic would not add any information related to the individual that was not known from the other information.

I have drawn the DfT's attention to the Upper Tribunal's recent decision in [NHS Business Services Authority v Information Commissioner and Spivack \[2021\] UKUT 192 \(AAC\)](#) – which dealt with identifiability, suppression of small numbers, and the application of section 40(2) – as well as the 2019 [ICO decision notice](#) in the same case.

I also highlighted the following in the ICO's [Code of Practice on Anonymisation](#) (page 31):

"Small numbers in small geographical areas present increased risk, but this does not mean that small numbers should always be removed automatically. For example, always removing numbers relating to five or 10 individuals or fewer may be a reasonable rule of thumb for minimising the risk of identification in a proactive disclosure scenario, but in the context of a specific freedom of information request a different approach may be possible, based on an application of the tests in the DPA."

The DfT's suppression of values between 1 and 4 has removed approximately 19% of all values from the disclosed table for Output Areas. That gap significantly reduces the analytic potential of the licensing data, for some purposes.

I am also concerned that neither the DfT nor the DVLA has provided a clear statement of their permission for me to re-use the disclosed information under the Open Government Licence. Both the DfT and the DVLA are Crown Copyright bodies and should, on the face of it, be able to grant that permission. However, currently I cannot document unambiguous permission to re-use the data under the OGL. This prevents me from distributing the data for open re-use by other parties.

Thank you for your attention to these concerns.

Yours faithfully,

Owen Boswarva