

# REPORT OF A BREACH OF THE CODE OF PRACTICE FOR STATISTICS

## 1. Core Information [\[guidance\]](#)

|   |   |
|---|---|
| <b>Title and link to statistical output</b>                       | <b>Unistats dataset (2019-20)</b><br><a href="https://www.hesa.ac.uk/support/tools-and-downloads/unistats">https://www.hesa.ac.uk/support/tools-and-downloads/unistats</a>  |
| <b>Name of producer organisation</b>                              | <b>HESA</b>   |
| <b>Name and contact details of person dealing with report</b>     | <b>Anna Sherratt, Principal Analyst, Office for Students.</b><br><a href="mailto:Anna.sherratt@officeforstudents.org.uk">Anna.sherratt@officeforstudents.org.uk</a> and<br><b>Rebecca Mantle, Official Statistics Manager, HESA</b><br><a href="mailto:Rebecca.Mantle@hesa.ac.uk">Rebecca.Mantle@hesa.ac.uk</a> |
| <b>Link to published statement about the breach (if relevant)</b> |   |
| <b>Date of breach report</b>                                      | <b>20/09/2019</b>   |

## 2. Circumstances of breach [\[guidance\]](#)

|   |   |
|---|---|
| <b>Relevant principle(s) and practice(s)</b>  | Principle T3 – Orderly release<br>Practice T3.4 – Statistical content has been made public or supplied to parties not essential for statistical production without prior permission for access, through accidental or wrongful release. |
| <b>Date of occurrence of breach</b>   | <b>10/09/2019</b>   |
| <p><u>Background</u></p> <p>The Unistats dataset contains information about courses offered by UK higher education providers. It is published in two ways – through a student-facing website and via the HESA website as a source data set. Since September 2018, the Unistats dataset has been published as an official statistic – originally by the Office for Students and then, this year, by HESA.</p> <p>The Office for Students manages the Unistats website on behalf of the other UK funders and regulators. The site is aimed at prospective students. The Unistats website presents the Unistats dataset in a user friendly and interactive format. On 12/09/2019, the Unistats website was superseded by a new website, called Discover Uni, also managed and delivered by the Office for Students on behalf of the other UK funders and regulators.</p> <p>HESA, as the designated data body for the Office for Students, publishes the Unistats source dataset set via its website. This is provided in XML and comma separated format under an Open Data licence and is aimed at more technically capable users, particularly those who wish to re-use the data for other purposes. The source data for Unistats undergoes a major revision annually, with minor revisions that may be as frequent as weekly between major updates.</p> <p>Until the launch of Discover Uni, HESA delivered a copy of the Unistats dataset to a website contractor working for the Office for Students to enable them to maintain the Unistats website. Due to the frequency of updates, this operational process was based upon use of an API to HESA systems, enabling the contractor to request the data as required through an automated process. HESA also shared the 2019 Unistats dataset with the website contractor delivering the new Discover Uni website.</p> |   |

### What happened

The 2019 version of the Unistats dataset was published by HESA on 11/09/2019 at 10.04 as an official statistic.

The Office for Students planned to publish the information from the 2019 Unistats dataset on the Discover Uni website on 12 September. However, the data was accidentally released onto the Unistats website at 20.40 on 10 September, preceding the official statistics release.

This happened because:

- (i) During the transition from one website contractor to another, HESA did not revoke the previous contractor's access to the API which enables the contractor to obtain the Unistats dataset.
- (ii) OFS did not communicate to this contractor that they should halt their weekly updates of the (shortly to be superseded) Unistats website in the run up to the publication of the official statistic.
- (iii) As a result, this contractor, working on behalf of the Office for Students, released the data onto the Unistats website.

### **3. Impact of the breach** [\[guidance\]](#)

There were around 1,500 views of the Unistats website between the accidental release of the data and the publication of the official statistic.

At this time of year, we believe that users are most likely to be looking for information about courses offered in the 2020-21 academic year. These users would have found what they were looking for. However, users who were looking for information about 2019-20 courses may have been misunderstood the information they were seeing (if they did not realise that the site had been updated), or else not found what they were looking for. We think that, on balance, this is unlikely to have had a material impact on users' decision making.

The contractor who received the dataset in error made no use of the data other than to present it on the Unistats website.

### **4. Corrective actions (taken or planned) to prevent re-occurrence**[\[guidance\]](#)

Historically the processes and structures in place to enable publication of this data were established when the organisation responsible for the primary website publication wasn't a producer of Official Statistics and the associated data transfers were treated by HESA as operational transfers alongside other routine data transfers that occur under authority vested by legislation. Both organisations involved in the publication of the data are now designated producers of Official Statistics and the processes and controls in place must fully reflect that.

1. HESA and the OfS have concluded that the early transfer of the Unistats dataset to a website contractor creates unnecessary risks because the desired timeliness of website publication can be achieved using a different business process. In future HESA will publish the Unistats dataset in compliance with all applicable Official Statistics requirements but will no longer provide early access to the website contractor through an automated process. Instead, HESA will deliver a public Application Programming Interface (API) that will make the dataset set available to the contractor at the point of general publication of the dataset. The website contractor will create an automated request to the API and launch the data within the website later on same the day the dataset is published by HESA. Occasionally, the contractor will require earlier access to the data for operational purposes. When this

is necessary, a representative of the Office for Students will convey this to HESA in writing, including the required date, time and destination of the transfer.

2. Office for Students standard contract terms have clear obligations on the supplier to hand over data, and to remove access to systems on contract termination. In addition, contract management best practice is currently being rolled out across the organisation. This includes detailed guidance around contract exit, including issues such as handover of data, intellectual property, removal of access to systems etc. In light of this breach, the Office for Students Procurement Team will re-prioritise the best practice rollout plan, and where there are contracts currently at risk of not being managed or exited effectively, immediately identify and train relevant contract managers. Additionally, the Procurement Team will review the existing guidance and contractual terms to ensure they remain fit for purpose.
3. HESA will ensure that all aspects of Unistats dataset dissemination are fully incorporated into their official statistics procedures and controlled by their official statistics team.
4. HESA and the Office for Students will review and strengthen communications between the OFS Head of Profession for Statistics and the HESA lead official for statistics.
5. The Office for Students will review communications between the Data, Foresight and Analysis directorate and the External Relations directorate, in particular to agree where responsibilities lie and to ensure that the programme management has the necessary input from data analysts.