

## The HMRC Datalab

### What is the HMRC Datalab?

The Datalab is a secure environment where independent institutions can access, free of charge, anonymised taxpayer and customs data, for research and analysis purposes.

### When was it set up?

This initiative was set up in May 2011.

### Why was it set up?

The Datalab is a key initiative of HMRC's Open Data Strategy, published in June 2012, which set out the ways in which HMRC is helping to deliver the Government's Transparency Agenda. By making the taxpayer-level data available in a safe way, the HMRC Datalab supports knowledge sharing and policy making, facilitating new areas of research using improved methodologies whose findings should benefit society.

### Who can use it?

Only approved researchers from trusted institutions can use this service. To become an 'approved academic researcher' researchers have to submit a proposal and complete a short training course, which includes legal issues as well as statistical disclosure control of outputs.

### How does it operate?

The operation of the Datalab is based on 5 key principles:

1. Safe environment: the Datalab computers are located in a secure room in Bush House on a standalone network that can only be accessed by authorised people
2. Safe data: information in the Datalab is anonymised so that no customer is directly identifiable
3. Safe projects: only research that serves one of HMRC's functions under the Commissioners for Revenue and Customs Act (2005) is allowed.
4. Safe researchers: only trusted researchers from trusted institutions are allowed to use the Datalab and have to undertake a training course before they can use the Datalab
5. Safe outputs: researchers are not allowed to take any information out of the Datalab until HMRC has checked that the outputs are non-disclosive, i.e. by aggregating up sufficiently to be so.

### How is it secure?

The Datalab follows HMRC data protection policies and there are restrictions on working practices to safeguard taxpayer confidentiality. As mentioned above, only 'approved academic researchers' can access this service. In order to be granted access to the Datalab, research must serve on HMRC functions under the Commissioner for Revenue and Customs Act 2005. Any output from analysis in the Datalab is subject to statistical disclosure control checks before release. In simple terms these aim to ensure that the outputs are sufficiently aggregated for no taxpayer's information to be directly or indirectly identifiable.

### Is the research independent?

Datalab research is not commissioned by HMRC. Research proposals are submitted by academic researchers working for themselves. HMRC simply use the Datalab facility to make taxpayer data available in a secure way.

## How does HMRC benefit from work in the Datalab?

Under current legislation, all Datalab research using HMRC data has to serve one of HMRC's functions (or be permitted under an existing legal gateway) – hence it should all contribute to HMRC's strategic objectives to a greater or lesser extent. This means HMRC can benefit from the projects. The current consultation includes a question on whether we should replace this legal requirement with the general condition that the research should deliver public benefits but need not necessarily benefit HMRC.