Office for National Statistics

Integrated household survey QMI

Contact:

socialsurveys@ons.gov.uk

Release date:
7 October 2014

Next release:
To be announced

Table of contents

1 . Methodology background

National Statistic	
Survey name	Integrated Household Survey (IHS)
Frequency	Annual
How compiled	Sample based survey
Geographic coverage	UK
Sample size	340,000
Last revised	7 October 2014

2 . Executive summary

The Integrated Household Survey (IHS) is a composite survey combining questions asked on a number of Office for National Statistics (ONS) social surveys to produce a dataset of "Core" variables. The aim of the IHS is the production of high-level estimates for particular themes to a higher precision and lower geographic level than previous ONS social surveys.

A set of "Core" questions were introduced to three survey modules in January 2008: the General Lifestyle Survey (GLF), Living Costs and Food Survey (LCF) and the Opinions Survey (OPN).

Subsequently, the English Housing Survey (EHS), Annual Population Survey (APS) and the Life Opportunities Survey (LOS) all joined the IHS core between 2008 and 2009.

Between 2010 and 2012, due to a variety of factors, the GLF, OPN, EHS and LOS modules were no longer included in the IHS to leave a dataset based upon the APS and LCF. Following a consultation in late 2013, a decision was made to remove the LCF from the IHS core from January 2014 for cost savings outweighing the benefits of inclusion.

From 2010 up until mid-2012, the IHS released annual data on a rolling quarterly basis, with a statistical bulletin released on an annual basis. However, following a consultation in late 2012 the decision was taken to reduce IHS to a single calendar year dataset as this was sufficient to meet user needs. The annual dataset for 2012, based upon APS and LCF data, was the first produced on this basis. The 2013 IHS consisted of a sample of around 190,000 addresses and an achieved sample of around 340,000 adults (representing around 0.5% of the UK population).

This report contains the following sections:

- Output quality

- About the output

- How the output is created

- Validation and quality assurance

- Concepts and definitions

- Other information, relating to quality trade-offs and user needs

- Sources for further information or advice

Output quality

This report provides a range of information that describes the quality of the output and details any points that should be noted when using the output.

We have developed [Guidelines for Measuring Statistical Quality](#); these are based upon the five European Statistical System (ESS) Quality Dimensions. This report addresses these quality dimensions and other important quality characteristics, which are:

- relevance

- timeliness and punctuality

- coherence and comparability

- accuracy

- output quality trade-offs

- assessment of user needs and perceptions

- accessibility and clarity

More information is provided about these quality dimensions in the following sections.

3 . About the output

Relevance

(The degree to which the statistical outputs meet users' needs.)

The Integrated Household Survey (IHS) has been developed by ONS to be a cost-effective way of obtaining a large dataset with a number of topics. It also enables analysis at lower level geographies that cannot typically be obtained with other social surveys. The IHS is used by academics, government departments and other stakeholders to provide evidence for planning, policy and monitoring purposes. It especially provides data on sexual identity, smoking prevalence and personal well-being to meet user needs.

Component surveys

In order to deliver its aim, the IHS created a core set of questions, which were to be asked across a series of ONS surveys. Due to a variety of factors, all but one survey (the Annual Population Survey) have dropped out of the IHS dataset since its inception. The IHS component surveys included:

- Annual Population Survey (APS): formed part of the IHS core between July 2008 and present; the APS comprised waves 1 and 5 of the Labour Force Survey (LFS) and an additional Local Labour Force Survey boost in England, Wales and Scotland, achieving a sample size of around 330,000 addresses per year

- Living Costs and Food Survey (LCF): formed part of IHS core between January 2008 and December 2013 (removed from IHS due to cost savings outweighing benefits of inclusion)

- General Lifestyle Survey (GLF): formed part of IHS core between January 2008 and December 2011 (survey stopped in December 2011)

- English Housing Survey (EHS): formed part of the IHS core between April 2008 and March 2011 (core questions made the survey too long)

- Life Opportunities Survey (LOS): formed part of the IHS core between June 2009 and December 2010 (could only be asked of wave 1 cross-sectional cases)

- Opinions and Lifestyle Survey (OPN): formed part of the IHS core between January 2008 and December 2009 (only asked of one household member so not appropriate for IHS)

Table 1 shows the surveys which have formed part of the IHS, and the dates of their involvement:

Table 1: IHS component surveys

	2008 Jan to Dec				2009 Jan to Dec				2010 Jan to Dec				2011 Jan to Dec				2012 Jan to Dec				2013 Jan to Dec			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
LFS or APS			APS: July 2008 to present																					
LCF			LCF: January 2008 to December 2013																					
GLF			GLF: January 2008 to December 2011																					
EHS				EHS: April 2008 to March 2011																				
LOS							LOS: July 2009 to Dec 2010																	
OPN	OPN: Jan 2008 to Dec 2009																							

Notes:

1. Q1 refers to January to March, Q2 refers to April to June, Q3 refers to July to September, Q4 refers to October to December.

The IHS collects a series of questionnaire topics, which are primarily driven from ONS Harmonised Standards. Topics include:

- basic descriptive information including: age, sex, marital status, and household or family relationships

- components of identity including: ethnicity, religion, national identity and sexual identity

- education, employment and economic activity including: educational attainment, employment status, and industry or occupation of employment

- perceived general health

- accommodation including: house type, length of residence and tenure

The IHS also collects some bespoke modules sponsored by government departments. These include:

- smoking prevalence rates

- veteran status

For more information about the IHS questionnaire, see the IHS user guidance.

Timeliness and punctuality

(Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.)

We published the first IHS data in September 2010, based on data from April 2009 to March 2010. This represented a 6-month lag between end of fieldwork and publication of results. A statistical bulletin accompanied this release of data, providing a summary of estimates for main topics of interest.

IHS data were published on a rolling quarterly basis, with an annual statistical bulletin, up to the release of April 2011 to March 2012 in September 2012. Subsequently, IHS data was published on an annual basis, accompanied with an annual statistical bulletin, as this met user demand.

The latest data available are the January to December 2014 data, which were published in early October 2015 (representing a 9-month lag between end of fieldwork and publication of results). This increased time lag was a result of changes to the procedures used to process data.

IHS data, together with supporting user guides, have been made available through the UK Data Service. Data and further information are also available from contacting ONS Social Surveys: socialsurveys@ons.gov.uk.

The large achieved IHS sample size enables users to conduct detailed analysis of question topics, enabling analysis at lower level geographies than is typically available on other surveys. Specifically, users are able to conduct analysis at local authority or unitary authority level for a large number of variables.

The datasets available from the IHS and achieved sample sizes for each IHS dataset, are presented in Table 2:

Table 2: IHS data and achieved sample

	Achived Sample (Adults)	Statistical Bulletin (publication date)
April 2009 to March 2010	449,330	September 2010
July 2009 to June 2010	442,957	
October 2009 to September 2010	433,410	
January 2010 to December 2010	424,816	
April 2010 to March 2011	420,695	September 2011
July 2010 to June 2011	399,164	
October 2010 to September 2011	374,218	
January 2011 to December 2011	361,229	
April 2011 to March 2012	349,576	September 2012
January 2012 to December 2012	338,174	October 2013
January 2013 to December 2013	342,186	October 2014

To date, the IHS has published six statistical bulletins, with the most recent bulletin published in October 2015, covering data collected between January to December 2014. Topics covered in this bulletin include sexual identity, perceived general health and smoking prevalence. Previous IHS bulletins have extended analysis to ethnicity, religion and occupation. Analysis in each bulletin is compared with main demographics including age, sex, geographic region, and other demographic characteristics.

For more details on related releases, the GOV.UK release calendar provides 12 months' advance notice of release dates. If there are any changes to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the Code of Practice for Official Statistics.

4 . How the output is created

Sampling

The Integrated Household Survey (IHS) sample frame is derived from its component surveys, which all use the Postcode Address File (PAF) of small users. This is a listing of all active address points maintained by the Post Office. The Small User file excludes addresses to which an average of 1,000 or more items of post are delivered per day. This reduces the likelihood of a business being selected in the sample.
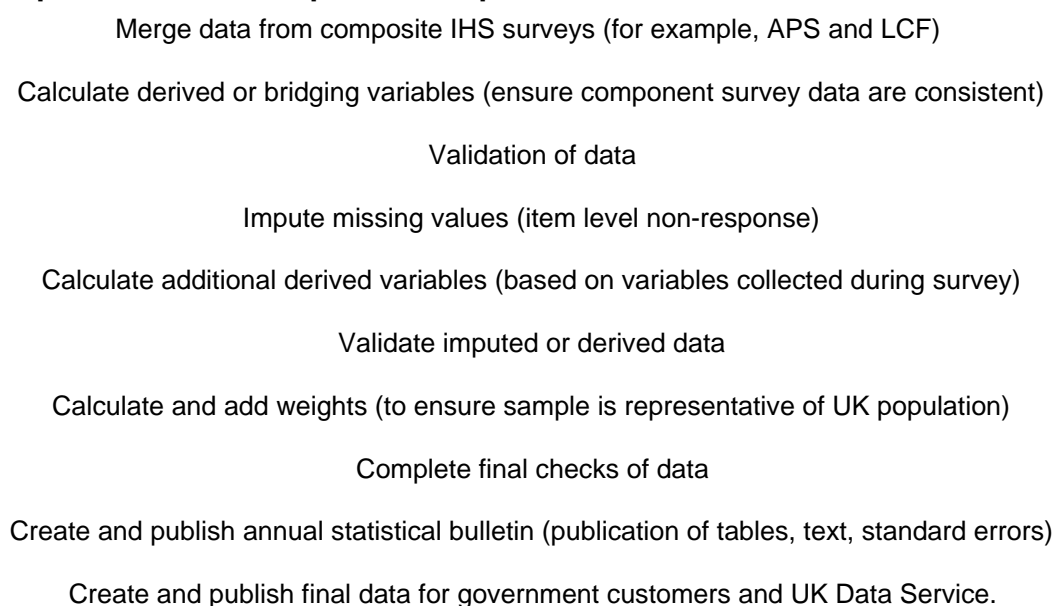
The sample design is also related to the constituent surveys of the IHS. The Annual Population Survey (APS) is an unclustered module which uses a single stage sample design. It therefore uses a sample of addresses selected directly from the PAF. The LCF is a clustered module which uses a multi-stage stratified random sampling design.

For more information about the IHS sample design, see the IHS user guidance.

Production processes

Once the data have been collected by the IHS component surveys, the process of creating, checking and validating the data takes place. Figure 1 shows the general steps involved in this process, based on the production of January to December 2013 data.

Figure 1: Steps involved in IHS production process

Merge data from composite IHS surveys (for example, APS and LCF)

Calculate derived or bridging variables (ensure component survey data are consistent)

Validation of data

Impute missing values (item level non-response)

Calculate additional derived variables (based on variables collected during survey)

Validate imputed or derived data

Calculate and add weights (to ensure sample is representative of UK population)

Complete final checks of data

Create and publish annual statistical bulletin (publication of tables, text, standard errors)

Create and publish final data for government customers and UK Data Service.

Information on the steps outlined in Figure 1 can be found in the IHS user guidance.

Proxy response

IHS questions are asked to all adults aged 16 and over in a sampled household. In certain circumstances, when a household member is unavailable for interview, interviewers accept information by proxy from another responsible adult in the household. The IHS survey allows proxy responses for all questions, apart from those relating to sexual identity.

Imputation

The IHS uses imputation to account and item non-response. This ensures the IHS data contains as much information as possible. Imputation takes place through a series of modules, which are specified by the core topics of interest, at the household level. Further information on imputation can be found in the IHS user guidance.

Weighting

Population weighting serves two purposes:

- it enables sample estimates to sum to known population estimates

- it compensates for differential non-response among different sub-groups in the population

Each sampled case is given a weight, which can be thought of as the number of people that case represents in the population. Ideally, each person in the population would have an equal probability of being selected for the IHS (as in a simple random sample). However, because of differential non- response, some people are more likely to be in the sample than others. People with a lower probability of being in the sample, such as young people in London, are assigned a higher weight.

The IHS uses a multi-stage population weighting procedure, which accounts for probability of selection and adjusts for non-response. Stages include:

- the initial address-level design weight

- an adjustment weight for multi-household addresses

- a non-response adjustment

- an attrition adjustment (where applicable)

- a scale factor for pooling surveys

- a calibration adjustment

Further information on weighting can be found in the IHS user guidance.

Statistical disclosure control

Statistical disclosure control methodology is applied to IHS data. The [Code of Practice for Official Statistics](#) sets out the principles for protecting data from being disclosed. More information can be found on the [ONS Disclosure Control](#) web page. This can cause some tables to have cells suppressed to avoid potential disclosure risk but detail and information is lost as a result.

5 . Concepts and definitions

(Concepts and definitions describe the legislation governing the output and a description of the classifications used in the output.)

The main definitions used in the IHS are explained in the [IHS user guidance](#). The IHS has adopted harmonised standards for question wording and guidance, which are driven by the ONS harmonisation programme. Information on these concepts and definitions can be found on the [harmonised standards](#) web page.

6 . Other information

Output quality trade-offs

(Trade-offs are the extent to which different dimensions of quality are balanced against each other.)

The publication of the IHS data and statistical bulletin takes into consideration the need to carry out rigorous quality assurance, and the need to publish data at the earliest opportunity to provide timely estimates for users. Further information on this can be seen in the "About the output" and "How the output is created" sections. However, the cost versus sampling accuracy has been a trade-off for the IHS.

Assessment of user needs and perceptions

(The processes for finding out about uses and users and their views on the statistical products.)

We regularly consult main stakeholders and users about the important decisions relating to the IHS. We chain an annual Steering Group, which provides an opportunity for main stakeholders to feed into the direction of the IHS. Information is also obtained from the UK Data Service regarding the use of IHS data, which is reviewed on an annual basis.

Finally, we have conducted a series of consultations with users about important decisions relating to the IHS. These include:

- [proposal to reduce the number of datasets produced using the Integrated Household Survey](#)

- [proposal to rotate the sexual identity questions on the IHS](#)

- [proposal to stop Integrated Household Survey core questions in the Living Costs and Food Survey](#)

7 . Sources for further information or advice

Accessibility and clarity

(Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.)

Our recommended format for accessible content is a combination of HTML webpages for narrative, charts and graphs, with data being provided in usable formats such as CSV and Excel. Our website also offers users the option to download the narrative in PDF format. In some instances other software may be used, or may be available on request. Available formats for content published on our website but not produced by us, or referenced on our website but stored elsewhere, may vary. For further information please refer to the contact details at the beginning of this report.

For information regarding conditions of access to data, please refer to the following links:

- Terms and conditions (for data on the website)

- Copyright and reuse of published data

- Pre-release access (ended from 1 July 2017)

- Accessibility

- Access to microdata via the Virtual Microdata Laboratory

In addition to this Quality and Methodology Information, basic quality information relevant to each release is available in the quality and methodology section of the relevant statistical bulletin.

User guidance for the IHS is also available.