Finding and using secondary data and resources for social science research

Beate Lichtwardt
UK Data Service/ UK Data Archive
University of Essex

Doctoral Open Day: Social Sciences 2
Who
... are we?

What
... do we offer?

Why
... could that be of interest to you?

Where
... to find data and resources?

How
... to access them?
Roadmap

- UK Data Service – overview service and data holdings
- Finding and accessing data
- Exploring data online
- User support and resources
Welcome to the UK Data Service

Your resource for quality social research data

A unified point of access to data from ESDS, Census Programme, Secure Data Service and others

Integrates ESDS, Survey Question Bank and Census.ac.uk

http://www.ukdataservice.ac.uk
What is the UK Data Service?

• a comprehensive resource funded by the ESRC

• a single point of access to a wide range of secondary social science data

• support, training and guidance

• listen to the recorded webinars at

http://ukdataservice.ac.uk/news-and-events/newsitem/?id=3494
Who is it for?

- academic researchers and students
- government analysts
- charities and foundations
- business consultants
- independent research centres
- think tanks

ukdataservice.ac.uk
What does the UK Data Service do?

- Put together a **collection of the most valuable data** and enhance that over time
- **Preserve data in the long term** for future research purposes
- Make the **data and documentation** available for reuse
- **Provide data management advice** for data creators
- **Provide support** for users of the service
- Information **about the use** to which data are put
- Easy access through a **website** - ukdataservice.ac.uk
What do users do with the data?

- Comparative research, restudy or follow-up study
- Re-analysis/secondary analysis
- Research design and methodological advancement
- Replication of published statistics
- Teaching and learning
Overview service – some statistics

Holdings: data for research and teaching purposes, used in all sectors and for many different disciplines

- **6,000** datasets in the collection
- **400** new datasets and new editions added within last 12 months
- **25,000** registered users
- **60,000** downloads worldwide p.a.
- **4000+** user support queries p.a.
Overview service – data sources

- Official agencies - mainly central government
- International statistical time series
- Individual academics - research grants
- Market research agencies
- Public records/historical sources
- Access to international data via links with other data archives worldwide
Types of data collections

- Survey microdata
  - cross-sectional
  - Panel/ Longitudinal
  - International microdata

- Aggregate statistics
  - International macrodata

- Census data
  - Aggregate data for 1971 - 2001
  - Microdata for 1991 and 2001 (2011 data are forthcoming)

- Qualitative and mixed methods data
  - in-depth interview transcripts, diaries, anthropological field notes, answers to open-ended survey questions, audio-visual recordings and images (mixed methods: combined with numeric data)
### Overview service – key data

<table>
<thead>
<tr>
<th>Government</th>
<th>Longitudinal</th>
<th>International</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large-scale government funded surveys</td>
<td>Major UK surveys following individuals over time</td>
<td>Multi-nation aggregate databanks and survey data</td>
</tr>
<tr>
<td>Census</td>
<td>Business</td>
<td>Qualitative</td>
</tr>
<tr>
<td>Census data 1971 - 2011</td>
<td>Sensitive data requiring secure access systems</td>
<td>Range of multi-media qualitative data sources</td>
</tr>
</tbody>
</table>
Key data

Key data

Browse our most popular datasets, by data type.

UK surveys  Cross-national surveys  Longitudinal studies  International macrodata  Census data

Business  Qualitative / mixed methods

Longitudinal studies involve repeated observations of the same subjects, allowing researchers to analyse change at the individual level.

The UK is home to several key longitudinal studies, including major panel and internationally-renowned cohort studies.

- Discover data from longitudinal studies
- Discover case studies relating to longitudinal studies

2011 CENSUS
The population census is a vital resource for social scientific research and policy development, providing a snapshot of demographic and social life in the UK that helps inform government
Large-scale UK surveys

- General Lifestyle Survey (General Household Survey)
- Labour Force Survey
- Health Survey for England/Wales/Scotland
- Living Costs and Food Survey (Expenditure and Food Survey)
- Crime Survey for England and Wales (British Crime Survey)
- Family Resources Survey
- Opinions and Lifestyle Survey (ONS Opinions Survey, ONS Omnibus Survey)
- English Housing Survey (Survey of English Housing)
- British Social Attitudes
- National Travel Survey
Benefits of the large-scale government data

• Good quality data
  • produced by experienced research organisations
  • usually nationally representative with large samples
  • good response rates
  • very well documented

• Continuous data
  • allows comparison over time
  • data is largely cross-sectional

• Hierarchical data
  • intra-household differences
  • household effects on individuals
Longitudinal data

- Longitudinal surveys involve repeated surveys of the same individuals at different points in time
- Allow researchers to analyse change at the individual level
- More complex to manage and analyse
Longitudinal data

Range of longitudinal data collections, but 9 ‘core’ collections in particular:

- National Child Development Study (NCDS)
- 1970 British Cohort Study (BCS70)
- Millennium Cohort Study (MCS)
- British Household Panel Survey (BHPS)
- Understanding Society (UKHLS)
- Families and Children Study (FACS)
- English Longitudinal Study of Ageing (ELSA)
- Longitudinal Study of Young people in England (LSYPE)
- Growing Up in Scotland (GUS)
The UK Household Longitudinal Study

- Started in 2009: UKHLS or ‘Understanding Society’; First data: 2010, 3 waves so far

- Study of the socio-economic circumstances and attitudes of 100,000 individuals in 40,000 British households

- UKLHS replaces and incorporates the BHPS (1991-2009, 5500 hh’s), retaining the latter’s longevity whilst adding
  - to the sample size
  - to the scope of the study (new components: innovation panel, greater detail on ethnic minority groups, qualitative and biomedical data collections)

- coverage includes:
  current employment and earnings; employment status; parenting and childcare arrangements; family networks; benefit payments; political party identification; household finances; environmental behaviours; consents to administrative data linkage (health and education).
British Birth Cohort Studies

- Impact of childhood conditions on later life and understanding children and families in the UK

- National Child Development Study follows a cohort born in a single week in 1958 - data collected at birth & ages 7, 11, 16, 23, 33, 41-42, 50 (7 Up TV series)

- 1970 British Cohort Study follows a cohort born in a single week in 1970 - data collected around birth & ages 5, 10, 16, 26, 29, 34, 38

- Millennium Cohort Study focuses on children born in 2000/2001 - sweeps at 9 months, 3, 6 and 8 years

- Wide range of social, economic, health, medical and psychological issues
International data

International data is split into two types of data:

- **Aggregate or macrodata**
  … data that has been aggregated to a country or regional level. This type of data is typically produced by inter-governmental organisations like the UN or World Bank. All the aggregate databanks contain time series data.

- **Survey or microdata**
  … data about individuals. This type of data is typically produced by an organisation in the academic sector and looks at values and attitudes.
International macrodata

- time series data aggregated to country/region
- regularly updated
- international governmental organisations (IMF, OECD, IEA, World Bank)
- wide range of socio-economic topics
- only available to staff and students from UK institutions of higher and further education (World Bank data is open access)
International data themes

Databanks cover:

- Economic performance and development
- Trade, industry and markets
- Employment
- Demography, migration and health
- Governance
- Human development
- Social expenditure
- Education
- Science and technology
- Land use and the environment
Example from UN COMTRADE, 2008

French snail imports

Graph: Celia Russell
International micro data

Data Service holds, and guides users towards accessing, several key international survey datasets, including:

- European Foundation for the Improvement of Living and Working Conditions datasets (Eurofund collection)
- European Social Survey (ESS)
- International Social Survey Programme (ISSP)
- European and World Values Surveys (EVS, WVS)
- European Election Study (EES)
- Young Lives (YL)
- Eurobarometers (EB); other regional Barometers such as: Afrobarometer, Arab Barometer, AsiaBarometer, Asian Barometer, New Baltic Barometer, New Democracies Barometer, New Europe Barometer, New Korea Barometer, New Russia Barometer
The World Values Survey, for example, reveals the extent to which people believe religion is important to them. In 36 out of 56 surveyed countries religion is 'very or rather important' to the majority of its people - regardless of the type of faith they believe in. In almost a quarter of the countries almost the entire population emphasises the prominence of religion (see graph below). At the lower end of the spectrum are Japan and China with about 20 per cent agreeing to religion being important compared to 100 per cent agreeing in Egypt and Jordan.

Data source: 2005-2008 World Values Survey
International data via other national archives

We can help users to locate and acquire data from other archives within Europe and worldwide, using a series of reciprocal agreements with the individual institutions.
Links with other data archives worldwide

CESSDA - Council of European Social Science Data Archives

Non-European: ICPSR, ASSDA
Qualitative data

The UK Data Service collections contain examples from a range of qualitative methods including:

- interview transcripts
- visual data
- focus groups
- essays
- diaries
- online data
- observation notes
- documents
- audio data
- newspaper clippings

etc.

Interview transcripts are the most commonly deposited qualitative data with the UK Data Archive.

Most contemporary data is deposited in a digital form, but we also hold (and sometimes digitise) legacy data which is a paper (handwritten or typewritten) format.
Qualitative data

Example


Teachers at a comprehensive school on the Isle of Sheppey were asked to set a particular essay to those pupils who were students in English lessons about ten days before they were due to leave school. The students were asked to imagine that they were nearing the end of their life, and that something had made them think back to the time when they left school. They were then asked to write an imaginary account of their life over the next 30 or 40 years.

The resulting data: 142 handwritten essays in 1978 by school leavers aged 15 and 16 years old.
Qualitative data

Example – How can the data be used?

Living and Working on Sheppey, Dawn Lyon and Graham Crow
• Digitised the original 1978 handwritten data.
• Coded and compared the data across time

Together, these two sets of essays shed light on the aspirations of Sheppey’s young people (and young people more generally) and cover a range of topics including health, education, career, family and leisure.

Example analysis of the career aspirations of Sheppey’s young people

“It was hard finding a job, I failed a few chances, but eventually got what I wanted locally, a craft apprenticeship” (Essay number 27, male)

“In my future I want to become either: a dance teacher, hairdresser, or a Professional Show Jumper/horse rider. If I do become a dancer my dream would be to dance for Beyoncé or someone really famous” (Essay number 61, female)
Qualitative data

Some popular datasets:

• Family Life and Work Experience Before 1918, 1870-1973

• Affluent Worker in the Class Structure: a Digitised Sample of the Luton Study, 1961-1962

• Health And Social Consequences Of The Foot And Mouth Disease Epidemic In North Cumbria, 2001-2003
UK census data

- 1971-2011 census data
- gold standard
- used as baseline for other statistics
- detailed combinations of characteristics
- small geographies
- census outputs
  - Aggregate data
  - Boundary data
  - Flow data
  - Microdata
- restricted to UK higher/further education staff and students, except for the microdata
Secure data

• Secure access to data that are too detailed, sensitive or confidential to be made available otherwise

• Data cannot be downloaded, researchers analyse the data remotely from their institutional desktop or in a safe room.

• We provide access to statistical and office software to make remote analysis and collaboration secure and convenient.
Secure data - examples

- Small Business Survey, 2010: Secure Data Service Access
- Millennium Cohort Study, 2001-2008, Linked Education Administrative Dataset: Secure Data Service Access
- Growing Up in Scotland: Sweep 5 Postcodes, 2009-2010: Secure Data Service Access
Finding and accessing data
How to search for data?

http://discover.ukdataservice.ac.uk/
How to search for data?

http://discover.ukdataservice.ac.uk/
Data access

• Web access to data and metadata via the UK Data Service web pages

• Documentation (codebooks, questionnaires) freely available to anyone browsing the web pages

• Data are freely available for use in higher education institutions

• Data supplied in a variety of formats
  - statistical package formats (e.g. SPSS, STATA)
  - databases and spreadsheets
  - word processed documents, PDF documents etc.
Accessing data - step by step

• Click on the ‘login’ link on the UK Data Service homepage
• Click on ‘login to the UK Data Service’
• Find your university in the list of institutions and proceed to the login page (or your UK Data Archive credentials if you are not at a UK University)
• Enter your login details
• Complete the registration form as a new user
• Accept the End User Licence
• Select the desired data from the Data Catalogue by clicking on the ‘Download/Order’ button
• Specify a project for which the data is to be used (minimum: 30 word description)
• Then:
  - download the data to local machine selecting your preferred format (SPSS, STATA, TAB etc.)
  or
  - place an order for the data (SC, if no online agreement; SL; AR) and complete all relevant forms
Data access conditions

• End User Licence (EUL)

• Special Conditions (SC)

• Special Licence (SL)
  - ONS - Approved Researcher Status
    (SL September 2005 - April 2008)
  - Other depositors - SL (e.g. ISER, CLS)

• Secure Data access
  - ONS-Approved Researcher or ESRC-Accredited Researcher
  - Training
  - Password and username
Example access conditions - UKHLS

SERIES ABSTRACT

The Understanding Society study, or the United Kingdom Household Longitudinal Study (UKHLS), which began in 2009, is conducted by the Institute for Social and Economic Research (ISER), at the University of Essex. As a multi-topic household survey, the purpose of Understanding Society is to understand social and economic change in Britain at the household and individual levels. It is anticipated that over time the study will permit examination of short- and long-term effects of social and economic change, including policy interventions, on the general wellbeing of the UK population. The Understanding Society study is a successor to the British Household Panel Survey (BHPS). The BHPS sample forms part of Understanding Society from Wave 2 onwards.

DATA ACCESS

GN 33423 | UNDERSTANDING SOCIETY: WAVES 1-, 2008- +

GN 33428 | UNDERSTANDING SOCIETY, 2008- : SPECIAL LICENCE ACCESS +

GN 33429 | UNDERSTANDING SOCIETY: WAVES 1-, 2009-: SECURE ACCESS -
Exploring data online
Online analysis with Nesstar

- Online data browsing and analysis system
- Allows users to search for, locate, browse and analyse and download a wide variety of statistical data within a web browser
- UK Data Archive, as service provider for the UK Data Service, hosts a Nesstar server populated by its most popular data series
- Registration is required for analysis such as crosstabulations and regressions

http://nesstar.esds.ac.uk/webview/index.jsp

<table>
<thead>
<tr>
<th>S4 MAIN Person Sex</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4 MAIN Feels about parents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not very good at being a parent</td>
<td>0.0</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>A person who has trouble being a parent</td>
<td>4.1</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td>An average parent</td>
<td>20.9</td>
<td>34.1</td>
<td>33.7</td>
</tr>
<tr>
<td>A better than average parent</td>
<td>30.4</td>
<td>27.3</td>
<td>27.4</td>
</tr>
<tr>
<td>A very good parent</td>
<td>41.5</td>
<td>34.3</td>
<td>34.5</td>
</tr>
<tr>
<td>Can't say</td>
<td>2.9</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>N</td>
<td>363.8</td>
<td>12,933.8</td>
<td>13,297.6</td>
</tr>
</tbody>
</table>

In general, frequencies and tables should be weighted before quoting percentages from them. To apply weights, select the Weight icon and choose one of the data.

Weight is on
International macrodata

Download the series you want from the appropriate online tool: Beyond 20:20, UKDS.Stat or IMF eLibrary.
International macrodata outputs

Proportion of seats held by women in national parliaments (%)

Mimas, University of Manchester. DOI: http://dx.doi.org/10.5257/wb/wdi/2013-01-30
Exploring census data

Explore census data

Census data can be accessed using a number of different tools.

- **Boundary Data Resources: Easy Download**
  Quick access to a set of the most frequently requested boundary datasets.

- **Boundary Data Selector**
  Download the boundaries you want, for the area you want, in the format you want.

- **Boundary Metadata Viewer**
  This facility allows users to search for and view metadata records associated with the census geographic datasets.

- **Casweb**

- **GeoConvert**
  GeoConvert is a geography matching and conversion tool. GeoConvert allows users to obtain and manipulate complex geographical and postcode data in a straightforward way.

- **InFuse**
  Access 2001 Census aggregate data for England and Wales by selecting
Religion

Percentage Christians, 2001

Percentage Christians, 2011

- More than 80%
- 70-79%
- 50-59%
- Less than 51%
Resources and help
User support and resources

- Help desk, individual user support
- Promotional events/ workshops
- Teaching data and resources
- User guides/ thematic guides
- Themes
- Online data analysis
- Advice on creating and managing data
- Geospatial resources
- Case studies
Teaching resources

- We can help provide support in many areas of teaching research methods:
  - Research design
  - Survey design and analysis
  - Software for analysis data
  - Secondary analysis of qualitative data
  - Managing research data
  - Consent and ethics

- Data collections are accompanied by user documentation and a record of main publications resulting from original use and re-use of the data

- Documentation should be recognised as a resource in its own right
Teaching with data

"Real data bring learning to life"

The use of data in teaching is an invaluable way for learners to engage with real-life research. By investigating the collections held at the UK Data Service, learners can find out how real data is originally conceived, collected, analysed and managed.

Data based on well-known studies or surveys can bring both substantive and methodological topics alive.

Using data in teaching can also enable critical engagement through the evaluation of the strengths and limitations of particular collections and their methodologies.

Our teaching datasets have been prepared for ease of use for learners. They are data files of a manageable size, which may have some features in a simplified form (e.g. weights). However, they are still real data which can be used to explore substantive
Teaching datasets in Discover

- **Discover**
  - Variable and question bank

- **Type**
  - Data collections (52)
    - Case studies (0)
    - Series (3)
    - ESRC outputs (0)
    - Support / how to guides (0)
    - Census data (2)

- **Subject**
  - Teaching data (52)
      - University of Manchester, Cathie Marsh Centre for Census and Survey Research. ESDS Government
      - Full record...
      - Explore online
      - Download/Order
      - Get full DDI XML
      - Similar data collections

    - SN 7210 Living Costs and Food Survey, 2010: Unrestricted Access Teaching Dataset
      - University of Manchester, Cathie Marsh Centre for Census and Survey Research. UK Data Service
      - Full record...
      - Explore online
      - Download/Order
      - Get full DDI XML
      - Similar data collections

    - SN 6968 Welsh Health Survey, 2010: Teaching Dataset (Adults)
      - National Centre for Social Research
      - Full record...
      - Explore online
      - Download/Order
      - Get full DDI XML
      - Similar data collections
Teaching resource: teaching sociology with archived data

"Incorporate research methods with analysing qualitative data"

This teaching resource enables students to develop a number of skills:

- students learn important research discovery skills including how to explore and search official websites
- they learn about the specific restrictions on certain data, as well as issues of confidentiality and data protection when working with personal data
- through the exploration of data collections they learn about a wide variety of research methods
Teaching resource: Interview methods

"Illustrating interview methods using our extensive data collection"

Interviewing is a frequently used method in social research with its suitability being entirely dependent on the particular research question. Qualitative interviewing is generally distinguished from questionnaire-based interviewing, even if the form of communication, such as face-to-face conversation, may be the same.

This teaching resource provides instructors and students with materials designed to assist in teaching qualitative interviewing.

The UK Data Service has created this resource to illustrate interview methods using its extensive data collection held as part of the UK Data Archive holdings, and, in particular, to assist instructors who have limited research materials of their own.
Help - Frequently Asked Questions (FAQ)

Answers to some of our most frequently asked questions are provided here. If you are unable to find an answer to your question please contact us by selecting one of the short web forms available from Help.

Questions about how to find and access data

How do I find a particular dataset?
How do I find studies on a particular subject?
Who can access the data?
Who can access census data?
How do I access the data?
What formats are the data available in?
How do I download/order data?
How much will it cost?
Contact

• Webpage for these slides:
  http://ukdataservice.ac.uk/news-and-events/eventsitem/?id=3666

Enquiries/ Help Desk:
  http://ukdataservice.ac.uk/help/get-in-touch.aspx

• Follow us on:
  https://twitter.com/UKDataService
  https://www.facebook.com/UKDataService
  https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=UKDATASERVICE
Questions?