The Irish Social Science Data Archive

Presentation to the 7th UL Winter School in Social Science Research Methods
Irisch Social Science Data Archive

**Background**

Founded in 2000 as a joint initiative of University College Dublin and the Economic and Social Research Institute, with the support of the Central Statistics Office.

Housed in UCD; originally in the Geary Institute; July 2012 moved to UCD Library.

**Objectives**

Provides a central access point for quantitative social science / public health data produced in Ireland; for research purposes.

Ensures wider access to datasets.

Facilitates secondary analysis.

Facilitates the wider acquisition of data analysis skills – teaching as well as research use.
Quantitative Data held by ISSDA

* **Quantitative**
  * Expresses data statistically or numerically

* **Examples**
  * The *number* of people who commit crimes each year
  * 30 *per cent* of voters regularly vote Fine Gael
  * The *rate* of suicide p.a. is one in every thousand

* **Microdata**
  * Data on individuals, households, business enterprises, discreet activities, etc.
  * Via surveys, observation, administrative records
  * Often aggregated to give larger picture
  * Produced by state agencies, polling agencies, organisations, researchers
  * Includes government studies: household surveys (e.g. Household Budget Survey), cohort studies, longitudinal studies (e.g. GUI, TILDA)
  * Otherwise mostly one-off studies, or in waves
  * Individual distribution (e.g. official website), by publication, and/or data archive such as ISSDA
Qualitative Data

* Data that are not numerical e.g.
  * Fieldnotes
  * Interviews
  * Focus Groups
  * Open-ended questions in surveys
  * Diaries
  * Direct observations
  * Film / audio / photographs

* Captures the *quality* of people’s behaviour. Qualitative data says something about the way people *experience* the social world and can be used to understand the *meanings* people give to behaviour

* Examples
  * Boyle (1977) studied the behaviour of a juvenile gang from the viewpoint of its members
  * Goffman (1968) tried to understand the experiences of patients in an American mental institution
Data Repositories / Archives

* Acquire datasets from producers
  * Maintain and build relationships with key data providers
  * Increasingly research grants are awarded on condition that data are deposited with an appropriate archive – part of the changing landscape that is “Open Access”
  * Anonymised

* Make datasets available to users
  * Promoting ISSDA and specific datasets
  * Presentations, webpages, twitter
  * Supporting users – application process, queries, referrals to data providers

* Preserve datasets for future use

**ISSDA = CONTENT, SERVICE AND PRESERVATION**
<table>
<thead>
<tr>
<th>Data</th>
<th>Primary data</th>
<th>Secondary data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>* Dataset collected by researcher for the specific</td>
<td>* Dataset collected by someone else for some other</td>
</tr>
<tr>
<td></td>
<td>purpose or analysis under consideration</td>
<td>purpose</td>
</tr>
<tr>
<td></td>
<td>* Research design and data collection processes are</td>
<td>* Researcher using secondary data for analysis</td>
</tr>
<tr>
<td></td>
<td>known</td>
<td>purposes did not participate in research design or</td>
</tr>
<tr>
<td></td>
<td>* Data collected to answer specific questions</td>
<td>data collection process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>* Data not collected to answer specific research</td>
</tr>
<tr>
<td></td>
<td></td>
<td>questions</td>
</tr>
</tbody>
</table>
Secondary Data Analysis: Advantages

- Use of existing data (e.g. from surveys and censuses) to test new ideas
  - Cost effective
  - Time saving
    - Can spend bulk of time analysing data
  - Breadth of data available
    - Many surveys available even in an Irish context
Secondary Data Analysis: Disadvantages

* Data not collected to answer your specific research questions
  * Therefore particular information you would like to have may not have been collected
  * May not have been collected in the geographic region you want to study
  * Or in the years you would have chosen
  * Or on the specific population that is the focus of your interest
  * Variables may have been defined or categorised differently than you would have chosen e.g. Race only as White / Other (rather than specific ethnicity)
Secondary Data Analysis: Disadvantages

- Data collection process: may not be known how it was done or how well it was done
- How seriously might the data be affected by a low response rate?
  - For both of the above check the methodological report
- Anonymisation of data may be quite extensive
- Check out the documentation of the data collection procedures, data cleaning procedures and other technical information
Locating Secondary Data

* Start with your research question and then seek the data?
* Check out the available secondary datasets and then formulate a research question?
* Example:
  * How does the experience of racism affect an individual’s health?
* Specify the population e.g. Children, adults, what ethnicities? National sample or restricted to a particular area?
* What variables? E.g. Race, age, gender, income, educational level
* What kind of data? E.g. Best addressed through a national survey, examination of hospital claims records, etc.
* Create a list of possible datasets that include information related to your research question...interactive process, may have to revise either your question or your data requirements depending on the data available to you
* Once dataset chosen examine the variables (what data might be missing); read documentation on data collection process, data cleaning procedures etc.
Irish Social Science Data Archive
The Irish Social Science Data Archive (ISSDA) is Ireland’s leading centre for quantitative data acquisition, preservation, and dissemination. Based at UCD Library, its mission is to ensure wide access to quantitative datasets in the social sciences, and to advance the promotion of international comparative studies of the Irish economy and Irish society.
Data

Application Procedures

Research/Educational Purposes

The data held by ISSDA can be ordered via this request form (MS Word), which is also downloadable on the webpages for the respective datasets. The form is an End User Licence, and users are legally bound by its terms. Please make sure that the form is fully completed and signed, otherwise, your application may be returned to you.

Completed EUL/application forms can be scanned and emailed to: issda@ocd.ie, alternatively, the forms can be posted to the ISSDA office:

Irish Social Science Data Archive (ISSDA),
UCD Library, Level 3,
University College Dublin,
Belfield,
Dublin 4,
Ireland

Teaching Purposes

Some of the datasets can be used for teaching purposes. Please check the dataset table below. There is a separate request form for teaching purposes, and you can find this on the individual dataset pages, where applicable. Please note that there may be additional delays with the application procedure for some of the datasets, as the data providers may need to be consulted. Please feel free to contact us with any queries.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Time Period</th>
<th>Data Provider</th>
<th>Research/Teaching</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growing Up in Ireland (GUI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave1 (9mo)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave2 (3yrs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave1 (9yrs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave1 (Time Use Data)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave2 (13yrs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Budget Survey (HSB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td></td>
<td>CSO</td>
<td>Apply to ISSDA</td>
<td>SPSS</td>
</tr>
<tr>
<td>1994/95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999/2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004/05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009/10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Datasets

* ISSDA main sources of data
  * CSO
  * ESRI
  * Major research projects
  * Commission for Energy Regulation

* CSO
  * Household Budget Survey
  * Quarterly National Household Survey
  * EU-SILC (Survey of Income & Living Conditions)
  * National Travel Survey

* ESRI
  * Growing Up in Ireland (GUI)
  * National Survey on Contraception and Pregnancy
  * National Survey of Housing Quality

* Major research projects
  * The All-Ireland Traveller Health Study
  * Children’s Sport Participation and Physical Activity
  * Survey on Lifestyle And Nutrition (SLáN)
  * The Irish Longitudinal Study on Aging (TILDA)
  * Doodle Den: Evaluation of the effectiveness of the Childhood Development Initiative’s Doodle Den Literacy Programme, 2008-2011
  * Etc.

* Commission for Energy Regulation
  * Smart Metering Project data (gas and electricity)
From Survey to Dataset

* **Dataset:** is the collective of all the questions that were asked in a survey and all the answers provided by all the respondents.

* **Variable Name:** survey questions are translated into variables and given a short name that is generally loosely based on the phrasing of the question. All the possible answers to each of the questions are allocated a numerical code.

* **Variable Label:** provides a more detailed description of the question than the Variable Name, as it can contain many more characters.

* **Codebook/Data Dictionary:** in order to be able to translate between the codes and the answers, a codebook is created. The codebook lists all the survey questions with the appropriate variable names and for each question it provides the different answer possibilities and the codes that were dedicated to each answer. Codes used to indicate non-response and missing data are also included. A data dictionary is a form of machine-readable codebook.
Depositing Studies to ISSDA

• **Dataset**
  - Can accept most formats – SPSS, SAS, Stata, CSV, etc.

• **Documentation**
  - Questionnaires / codebooks (in-depth explanation of questionnaire, plus summary statistics)
  - Technical information – weights, sampling, constructed variables, regional variables

• **Anonymisation**
  - Data providers remove direct identifier variables – names, addresses
  - Data providers may need to recode or combine variables in cases of small numbers (e.g. ages)
DEPOSIT DATA

WHY DEPOSIT?
Sharing your data via ISSDA ensures that your data will be professionally curated, will be easily accessible to users now and in the future, and will help to increase the impact and visibility of your own research.

More information on the importance of data sharing is available from UCD's Research Data Management guide: http://libguides.ucd.ie/data-sharing

WHO CAN DEPOSIT?
We acquire data from academic, research bodies and public sector sources, supporting:
- Archival preservation
- Secondary use and analysis for research
- Teaching and learning use
- Replication and validation of research

We are happy to discuss any offers of data that come within the thematic scope of the ISSDA collections, broadly, these relate to Irish society and include the societal aspects of environmental and medical data. Please see ISSDA's Collection Development Policy for an outline of the scope of our collections and criteria for evaluating datasets.

HOW TO DEPOSIT

1. Preparing data for deposit
Depositing Studies to ISSDA

Please give information about the response rate. Give the number of valid observations in the database and the non-response (at each stage). A flow-chart diagram would be useful here:

**Geographical Coverage**
- Country: Ireland
- Region: Dublin
- County: Dublin
- Town: Tallaght West
- Administrative unit: 7 schools
- Other geography:

**Methods of data collection**
Select from list or provide free text using the “other” box:
- **✓** PAPI (Paper and Pencil Interviewing):
  - With interviewer (face to face, telephone or other)
- **✓** PAPI (Paper and Pencil Interviewing):
  - Self-completion (distributed by post, email or other)
Prior to deposit to ISSDA all data studies are checked by CSTAR and a report submitted on:

* General observations
* Suitability for deposit to ISSDA (e.g. data quality, anonymisation, provision of contextual / descriptive information for users)
* Specific actions to address the above, where necessary
Growing Up in Ireland

Study documentation

- Guides (PDF):
  - Wave 1: Infant Cohort
  - Wave 1: Child Cohort
  - Wave 1: Child Cohort - Time Use Data
  - Wave 2: Infant Cohort

- Data dictionaries (PDF): Wave 1: Child Cohort - Wave 1: Infant Cohort - Wave 2: Infant Cohort

- Questionnaires (PDF): Wave 1: Child Cohort - Wave 1: Infant Cohort - Wave 2: Infant Cohort

- Sample designs (PDF): Wave 1: Child Cohort Design Sample - Wave 1: Infant Cohort

- Variable naming and longitudinal data dictionary (PDF): Wave 1 & 2: Infant Cohort
GUI Summary Guide

- Background to the study
- Sample design (for each cohort)
- Instrument development (design and piloting)
- Fieldwork and implementation (how it was carried out)
- Structure and content of the data files (e.g. main respondent, variable names, scaled measures used)
- Ethical considerations
G. CHILD’S ACTIVITIES

Now I would like to ask you about some of the Study Child’s day-to-day activities.

G1. [Show Card G1] Looking at Card G1, on a normal weekday during term time, how many hours does the Study Child spend watching television, videos or DVDs? Please remember to include time before school as well as time after school?

None ........................................... 1 3 hours to less than 5 hours ................. 4
Less than an hour ............................ 2 5 hours to less than 7 hours ................. 6
1 hour to less than 3 hours .................. 3 7 hours or more ................................. 8

G2. [Show Card G2] Looking at Card G2, on a normal weekday during term time, about how many hours does the Study Child spend reading for pleasure [NOT during school hours]? Include time when the child reads to themselves or is read to by someone else. Do not include time spent listening to books on audio tapes, records, cds or a computer.

None ........................................... 1 5 hours to less than 7 hours ................. 4
Less than an hour ............................ 2 7 hours or more ................................. 8
1 hour to less than 3 hours .................. 3 Child can’t read .................................. 7
3 hours to less than 5 hours .................. 4

G3. [Show Card G3] Looking at Card G3, on a normal weekday, during term-time, about how much time does the Study Child spend using the computer. Please include time before school as well as time after school. DO NOT include time spent using computers in school.

None ........................................... 1 3 hours to less than 5 hours ................. 4
Less than an hour ............................ 2 5 hours to less than 7 hours ................. 6
1 hour to less than 3 hours .................. 3 7 hours or more ................................. 8
GUI Data Dictionary (Codebook)

* Provides information on every variable in the dataset, including:
  * Label
  * Type of variable
  * Measurement level
  * Valid values
### MMA4

<table>
<thead>
<tr>
<th>Standard Attributes</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>8</td>
</tr>
<tr>
<td>Label</td>
<td>A4. How many people in household</td>
</tr>
<tr>
<td>Type</td>
<td>Numeric</td>
</tr>
<tr>
<td>Measurement</td>
<td>Nominal</td>
</tr>
<tr>
<td>Valid Values</td>
<td>One</td>
</tr>
<tr>
<td></td>
<td>Two</td>
</tr>
<tr>
<td></td>
<td>Three</td>
</tr>
<tr>
<td></td>
<td>Four</td>
</tr>
<tr>
<td></td>
<td>Five</td>
</tr>
<tr>
<td></td>
<td>Six</td>
</tr>
<tr>
<td></td>
<td>Seven or more</td>
</tr>
</tbody>
</table>

### mma5ap1

<table>
<thead>
<tr>
<th>Standard Attributes</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>9</td>
</tr>
<tr>
<td>Label</td>
<td>Gender P1</td>
</tr>
<tr>
<td>Type</td>
<td>Numeric</td>
</tr>
<tr>
<td>Measurement</td>
<td>Nominal</td>
</tr>
<tr>
<td>Valid Values</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Female</td>
</tr>
<tr>
<td>Missing Values</td>
<td>Refusal</td>
</tr>
<tr>
<td></td>
<td>Don’t know</td>
</tr>
<tr>
<td>TopicCat</td>
<td>Varname</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>HC</td>
<td>pH26a</td>
</tr>
<tr>
<td>HC</td>
<td>pH26b</td>
</tr>
<tr>
<td>HC</td>
<td>pH26c</td>
</tr>
<tr>
<td>HC</td>
<td>pH27a</td>
</tr>
<tr>
<td>HC</td>
<td>pH27b</td>
</tr>
<tr>
<td>HC</td>
<td>pH28a</td>
</tr>
<tr>
<td>HC</td>
<td>pH28b</td>
</tr>
<tr>
<td>HC</td>
<td>pH30</td>
</tr>
<tr>
<td>HC</td>
<td>sHC30</td>
</tr>
<tr>
<td>HC</td>
<td>pH31a</td>
</tr>
<tr>
<td>HC</td>
<td>sHC31a</td>
</tr>
<tr>
<td>HC</td>
<td>pH31b</td>
</tr>
<tr>
<td>HC</td>
<td>sHC31b</td>
</tr>
<tr>
<td>HC</td>
<td>pH31c</td>
</tr>
<tr>
<td>HC</td>
<td>sHC31c</td>
</tr>
<tr>
<td>Name</td>
<td>Type</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>bpc2C8c</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C9a</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C8d</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C9b</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C10</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C11</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C12</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C13</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C14</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C15</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C16a</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C16b</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C16c</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C16d</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C16e</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C16f</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C16g</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C17</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C18a</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C18b</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C18c</td>
<td>Numeric</td>
</tr>
<tr>
<td>bpc2C19d</td>
<td>Numeric</td>
</tr>
</tbody>
</table>
In the News…

Better-off women more likely to drink alcohol in pregnancy

Growing Up in Ireland study. Data reveals one in 10 smoked while expecting a baby.

Women with higher levels of income and education are more likely to drink alcohol weekly during their pregnancy, new figures have shown.

While more than one in 10 mothers (13 per cent) smoked all the way through their pregnancy, there was an increased risk of smoking linked to being poor and having lower levels of education.

The findings are from a new report, Maternal Health Behaviours and Child Growth in Infancy, published as part of the Growing Up in Ireland study.
Selecting a Dataset

- Increasing amount of information on webpage:
Accessing Datasets

* Application forms and end-user licenses available on website
* Received by post or as scan to ISSDA
* Details checked
* Dataset supplied by secure email (HEAnet’s FileSender)

**TERMS OF USE**

4. To ensure that the means of access to the data collections (such as passwords) are kept secure and not disclosed to a third party (excluding those listed in Appendix 1) except by special written permission or licence obtained from the ISSDA.

5. Not to use the data collections to attempt to obtain or derive information relating specifically to an identifiable individual or household.

6. To be aware at all times of the risk of inadvertently disclosing information, which might result in the identification of an individual. All use of the data and production of all analysis and output should be sensitive to this risk.


8. To acknowledge, in any work based in whole or part on resources provided by the ISSDA, the original data creators, depositors or copyright holders and the ISSDA, and to declare, in any such work, that those who carried out the original analysis and collection of the data bear no responsibility for the further analysis or interpretation of it.
Queries

* General information on how to access data

* “In the verbal fluency task, data is presented that indicates how many animals were named in one minute. I wonder if data is available which gives a breakdown of the actual animals recalled, in what order and what timeframe within the one minute” (TILDA)

* “Please can I access the following variables relating to measured BMI...” (SLAN)

* “When will the second round of the child cohort data be made available?” (GUI)
FAQs: General

1. Access

**Obtaining Data:**

**Q. Who is permitted to access data through ISSDA?**

A. Any student or researcher from any organisation, country or field of study can apply to ISSDA to access our datasets. Access to some datasets may be restricted in some cases, for example:

  **Teaching Purposes:** To check which datasets can be used for teaching purposes, please refer to ISSDA’s data page (www.ucd.ie/issda/data) where the “Purpose” column will indicate this. If the “Purpose” column indicates that the dataset can be used for teaching purposes please fill out a teaching request form located on the ISSDA website. The license is for once-off use of data. Should you wish to use the data again when re-running the module, please submit a new request form.

  **Commercial Purposes:** To check if you can use the datasets for commercial purposes please contact us at issda@ucd.ie and we will contact the data provider to check on your behalf.

**Q. Is there a charge to access data through ISSDA?**

A. Accessing data through ISSDA is a free service and there is no charge.

**Forms:**

**Q. How do I obtain the dataset?**

A. In order to access the data you need to first submit an application form which you can access through the following link: http://www.ucd.ie/issda/data/ (the link to the request form is located at the top of the web page). The application form is also accessible through the webpages of individual datasets. The form needs to be signed by you (the End User) as it is an End User Licence Agreement. You can then scan the application form or attach a jpeg of your signature and email it to us at issda@ucd.ie or post it to us at Irish...
FAQs: GUI

Using the data

Q. There is so much information in the Anonymised Microdata File - how do I get started using the data?

A. The first step is to look through the paper versions the questionnaires which are available to download from the Growing Up in Ireland website http://www.growingup.ie/index.php?id=7. These are much easier to navigate than the datafile – they are divided into sections by topic and clearly show the routing of the questions.

Q. There are two versions of the file provided for the Infant Cohort, with different variable naming conventions – Convention A and Convention B – which one should I use?

A. Variable naming convention A was developed at Wave 1 and the variable names link back to the paper questionnaires. You can use this convention if you know you are only interested in doing analysis on one wave of data and do not intend to match the two waves.

If you do intend to match the two waves of data to do cross-wave analysis, variable naming convention B is better suited to this type of analysis. Variable names are not linked to the questionnaires. Any variables which are asked in both waves will have the same core variable name with wave one prefixed with 'a' and wave two prefixed with 'b'.

ISSD
Irish Social Science Data Service
ISSDA’s Users

* Researchers working on academic, non-profit, and consultancy projects
  * Masters, PhDs, Postdocs, academics

* Undergraduates (e.g. economics, psychology, sociology)

* Government departments/agencies

* International demand for Irish data
  * UK, USA and other EU countries

* Discipline areas – wide
  * Social sciences, business/commerce, psychology, public health, engineering, agriculture, etc.
* Examination of the impact of working arrangements on family life in Ireland
* Exploration of rural ageing in Ireland today. The focus will be on the social, health and economic lives of older adults living in different kinds of rural settings and where appropriate different kinds of residential environments.
* Research on the social stratification of early childhood care and education
* Data will be used to examine the relationship between social advantage and involvement in sport and physical activity amongst disadvantaged groups in Ireland
* The data is to be used for the analysis of market trends and familial change before and during the economic crisis
* I intend to use the GUI data to analyse children's attitudes to Irish language learning in order to contextualise my PhD study which aims to develop a new approach to teaching the Irish language.
We want to analyse how household spending responds to actual and anticipated labour market and income shocks. This involves constructing a pseudo panel dataset from the various SILC waves to analyse how similar households respond to shocks over time.

I plan to use the HBS data along with price indices from the CSO to calculate a set of inflation rates for different subsets of the population and to carry out inequality analysis also. The end date of the project would be 2016.

QNHS data will be used for (a) research project on personal mobility and the economic crisis (funded by internal TCD funds) and (b) changing employment and migration since 2006 (continuation with internal TCD funds of project funded 2013-14 by the Irish Research Council).
OTHER ARCHIVES, PORTALS AND SOURCES
The Irish Qualitative Data Archive (IQDA) is a central access point for qualitative social science data generated in or about Ireland. The archive frames the parameters and standards for archiving qualitative data within the Irish research community. The IQDA is a consortium member in the Digital Repository of Ireland.
Welcome to data.gov.ie!

The departments of the Government and other public sector bodies are releasing public data to help people understand how government works, how decisions are made, and what we know about the country and society. This portal brings these datasets together in a single searchable website. Learn more...

What data can be found here?

Data.gov.ie is a searchable directory of datasets that are published on the websites of Irish public sector bodies.

The portal was initially populated through a review of the websites of government departments and state bodies. This initial audit did not include the websites of local authorities and state-sponsored bodies (An Post, RTE, Bord na Móna, etc.).

Datasets are listed on the portal if they are available in a machine-processable form that allows re-use and re-processing of the data. This excludes datasets that are only available in PDF and Word format.
Additional Irish Sources

* Ageing in Ireland [www.agestats.ie/](http://www.agestats.ie/)

* Public health [www.thehealthwell.info/](http://www.thehealthwell.info/)

* Health [www.hiqa.ie/healthcare/health-information/data-collections/online-catalogue](http://www.hiqa.ie/healthcare/health-information/data-collections/online-catalogue)

Users from outside the UK can register and access most data in the Data Catalogue.

To register, a username and password are required.

Due to the data redistribution licence agreements with dataset providers/service funders, registration will not permit non-UK users access to a small number of datasets.
An umbrella organisation for social science data archives across Europe.

The CESSDA Catalogue enables users to locate datasets, as well as questions or variables within datasets, stored at archives throughout Europe.

Data collections include sociological surveys, election studies, longitudinal studies, opinion polls, and census data from the European Social Survey, the Eurobarometers, and the International Social Survey Programme among others.
CESSDA Keyword Search on Age Discrimination

Search Term: AGEDISCRIMINATION
Inter-university Consortium for Political and Social Research (ICPSR)

- World’s largest archive of digital social science data
  - www.icpsr.umich.edu/

* ISSDA subscribes for all RoI Universities

* Create a MyData account to be able to download data as an ICPSR member
Search Examples

* By topic e.g. income health china

* By variable e.g. read newspapers, political party, volunteer; body mass, difficulty walking

* By specific research question (using natural language) e.g. Does health insurance change the use of medical services? Or broken down by concept e.g. income health china

* By specific dataset e.g. "World Values Survey"
Variable Search Example

Variable Search Results

Use the search box below to revise your search, or start a new search.

read newspapers, political party, volunteer

1st secondary activity

At any time while you were [REPEAT ACTIVITY] did you do anything else? (like talking, reading, watching TV, listening to the radio, eating or caring for children) ...

(view details)

Taken from: National Survey of Parents, 1999-2001 - Time Diary Data.

2nd secondary activity

At any time while you were [REPEAT ACTIVITY] did you do anything else? (like talking, reading, watching TV, listening to the radio, eating or caring for children) ...

(view details)

Taken from: National Survey of Parents, 1999-2001 - Time Diary Data.
IPUMS-International is a project dedicated to collecting and distributing census data from around the world. Its goals are to:

- Collect and preserve data and documentation
- Harmonise data
- Disseminate the data absolutely free!

Ireland (C.S.O.) has contributed 9 census samples

### Popular Variables

#### Top Ten
- Educational attainment
- Marital status
- Employment status
- Relationship to head
- Age
- Sex
- Class of worker
- Years of schooling
- Ownership of dwelling
- Urban-rural status

#### 11 to 20
- School attendance
- Occupation, ISCO
- Literacy
- Industry, general recode
- Country of birth
- Nativity status
- Children ever born
- Occupation, unrecoded
- Industry, unrecoded
- Religion

#### 21 to 30
- Water supply
- Children surviving
- Electricity
- Sewage
- Migration status, 5 years
- Earned income
- Total income
- Spouse's location in hh
- Number of rooms
- Toilet
# Data Harmonisation

## Marital Status: IPUMS-International

<table>
<thead>
<tr>
<th>Bangladesh 2011</th>
<th>Mexico 1970</th>
<th>Kenya 1999</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = Unmarried</td>
<td>1 = Married, civil &amp; relig</td>
<td>1 = Never married</td>
</tr>
<tr>
<td>2 = Married</td>
<td>2 = Married, civil</td>
<td>2 = Monogamous</td>
</tr>
<tr>
<td>3 = Widowed</td>
<td>3 = Married, religious</td>
<td>3 = Polygamous</td>
</tr>
<tr>
<td>4 = Divorced/separated</td>
<td>4 = Consensual union</td>
<td>4 = Widowed</td>
</tr>
<tr>
<td></td>
<td>5 = Widowed</td>
<td>5 = Divorced</td>
</tr>
<tr>
<td></td>
<td>6 = Divorced</td>
<td>6 = Separated</td>
</tr>
<tr>
<td></td>
<td>7 = Separated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8 = Single</td>
<td></td>
</tr>
</tbody>
</table>
Recent papers using IPUMS-I data

- Occupational Segregation of Afro-Latinos
- Cross-National Comparison of Internal Migration Age Profiles
- Changing Household Patterns of Young Couples in Low- and Middle-Income Countries
- Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure
- Local Economic Conditions and Participation in the Rwandan Genocide
International Data Providers

* WHO: [www.who.int/research/en/](http://www.who.int/research/en/)


Find Useful Data and Statistics

Google Data Search

Why trawl the internet for useful data sources, when the SSL librarians have done it for you? Use the Google Custom Search box below to find relevant data from over 1000 carefully selected online statistical sources.

[Google Custom Search]
Using Google to Locate Data

- children ireland dataset OR "data archive" OR "data portal"

- Gender dataset OR "data archive" OR "data portal"

- migration dataset OR "data archive" OR "data portal"
Contact Details

ISSDA: www.issda.ie

Email: issda@ucd.ie

Julia Barrett, Research Services
Manager, UCD Library

Julia.barrett@ucd.ie

See also the new Research Data Management Guide at:

http://Libguides.ucd.ie/data

Ireland’s centre for quantitative data acquisition, preservation, and dissemination.

Established in 2000, ISSDA’s mission is to ensure wide access to quantitative datasets in the social sciences, and to advance the promotion of international comparative studies of the Irish economy and Irish society. To this end it acquires data from academic, research bodies and public sector sources, supporting:

• Secondary use and analysis for research
• Teaching and learning use
• Replication and validation of research
• Archival preservation

At the broadest level, ISSDA collects anonymised microdata, i.e. data that arises from surveys or administrative records relating to individuals or groups of individuals that have been cleansed of personal identifiers or other information that may enable identification of individuals represented in a dataset. Datasets collected by ISSDA generally have an Irish geographic coverage.

Access A Wide Variety Of Key Datasets In The Social Sciences!

While based in UCD and managed by UCD Library, ISSDA is a national service that provides access to a wide range of data. Information about the data is available from

Other ISSDA Services

Inter-university Consortium for Political and Social Research (ICPSR)

ISSDA holds a national subscription to the ICPSR data archive. This means that if you are based in any Irish university (DCU, NUI, NUIM, TCD, UCC, UCD, UL) you will have access to this data archive. The ICPSR is an international consortium of approximately 700 academic institutions and research organisations that maintains an archive of more than 500,000 data files in the political and social sciences. To access the ICPSR database, you will need to set up a MyData account, or alternatively log on with a Facebook or Gmail account:

https://www.icpsr.umich.edu/rpLogin

Training and Support