

Finding & accessing data

Jenny O'Neill, Data Manager

Overview

- ◎ Introduction to ISSDA

- ◎ 1. Key concepts
- ◎ 2. From survey to dataset
- ◎ 3. Finding data
 - Ireland
 - Europe
 - International
- ◎ 4. Accessing data



“

ISSDA
is Ireland's leading centre for
quantitative
data **acquisition, preservation,
and dissemination.**

ISSDA Objectives

**Central
access
point**

**Facilitates
secondary
analysis**

**Wider
access to
datasets**

**Acquisition
of data
analysis
skills**



1.

Key concepts

Primary data

- ⦿ Data collected by a researcher for the specific research purposes
- ⦿ CSO
- ⦿ ESRI
- ⦿ Major research groups
- ⦿ Individual researchers

D A T A S H A R I N G

Secondary data

- ⦿ Dataset collected by someone else for some other purpose
- ⦿ Researcher using secondary data for analysis purposes did not participate in research design or data collection process
- ⦿ Data not collected to answer specific research questions

Quantitative Data

Expresses data
statistically or
numerically



Quantitative Data held by ISSDA

- ◎ 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 - captures
the quality of people's
behaviour



Qualitative Data

- ⊙ Fieldnotes
- ⊙ Interviews
- ⊙ Focus Groups
- ⊙ Open-ended questions in surveys
- ⊙ Diaries
- ⊙ Direct observations
- ⊙ Film / audio / photographs

Example

- ⊙ Gray, Jane, O'Carroll, Aileen, Ó Riain, Seán, & Geraghty, Ruth. (2015) Life Histories and Social Change collection, Digital Repository of Ireland [Distributor], Irish Qualitative Data Archive [Depositing Institution], DOI: 10.7486/DRI.9593xp97w

Types of studies

One off
e.g. All Ireland
Traveller Health
Survey, Doodle Den
Literacy Programme

Repeated cross-
sectional
e.g. Healthy Ireland,
Irish Sports Monitor

Longitudinal
e.g. TILDA, Growing Up
in Ireland, Preparing
for Life



Data Repositories & Archives



Acquire datasets from Data Providers. Maintain and build relationships with key Data Providers.



Increasingly research grants are awarded on condition that data are deposited with an appropriate.



Data are anonymised before deposit



Make data available via a Data Catalogue or website



Promote services via presentations, webpages, twitter



Supporting users – application process, queries, referrals to data providers



2.

From survey to dataset

Datasets



ISSDA data

- ⊙ CSO
- ⊙ ESRI
- ⊙ Major research projects
- ⊙ Commission for Energy Regulation



CSO

- ⊙ HBS
- ⊙ QNHS
- ⊙ EU-SILC
- ⊙ National Travel Survey



ESRI

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



Major research projects



- ⊙ Healthy Ireland
- ⊙ Irish Sports Monitor
- ⊙ The Irish Longitudinal Study on Aging (TILDA)
- ⊙ Irish Student Engagement Survey (ISSE)



CER

- ⊙ 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

- ⦿ [File Format Policy](#) – SPSS, SAS, Stata

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)

Depositor Forms

- ⦿ License agreement
- ⦿ Study Metadata

Study level metadata

⦿ **Responsible Parties**

⦿ **Collection Description**

⦿ Title of the Data Collection

⦿ Population

⦿ Temporal coverage

⦿ Methods of data collection

Description or Abstract

Observation units

Geographical Coverage

⦿ **Data and Documentation: Files' Description**

⦿ **Details of Methodology and Sampling Procedures**

Variable level metadata

- ⊙ Full question text
- ⊙ Universe
- ⊙ Interviewer instructions

Anonymised data

AMF

Anonymised
Microdata File

VS.

RMF

Researcher
Microdata File

3.

Finding data

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?

Locating Secondary Data

- ◎ 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.



Economic and Social Research Institute

The ESRI is one of Europe's leading research centres in the social sciences, focusing on economic and social change in Ireland.

CSO
ESRI



Growing Up
ICPSR
SLÁN

News

Teagasc National Farm Survey (2000-2015) now available
Harmonized TILDA now available
September Closures

Help

Visit our Help section for:

- [FAQs](#) relating to access, receipt of the datasets, and analysis;
- [Information Flyers](#) about ISSDA;
- [links to Other Data Archives](#);
- and [Researcher and Teaching Resources](#).

Accessing Data

3 Simple Steps to Apply for Data

1. Identify the dataset(s) you require.
2. Complete the [request form for research purposes](#) or [request form for teaching purposes](#). (This form is requested by the data providers.)
3. Send the **signed** consent form by [email](#), or to: ISSDA, UCD Library, UCD, Belfield, Dublin 4.
Data can be delivered online or via post. If you have any queries regarding the above, please contact us at issda@ucd.ie

Selecting a Dataset

Increasing amount of information (metadata) on webpage:

◎ Healthy Ireland:

<http://www.ucd.ie/issda/healthyireland/>

◎ Irish Sports Monitor:

<http://www.ucd.ie/issda/irishsportsmonitor/>

◎ Doodle Den: <http://www.ucd.ie/issda/doodleden/>

Other Archives & Data Sources

Ireland

IQDA

Irish Qualitative Data
Archive

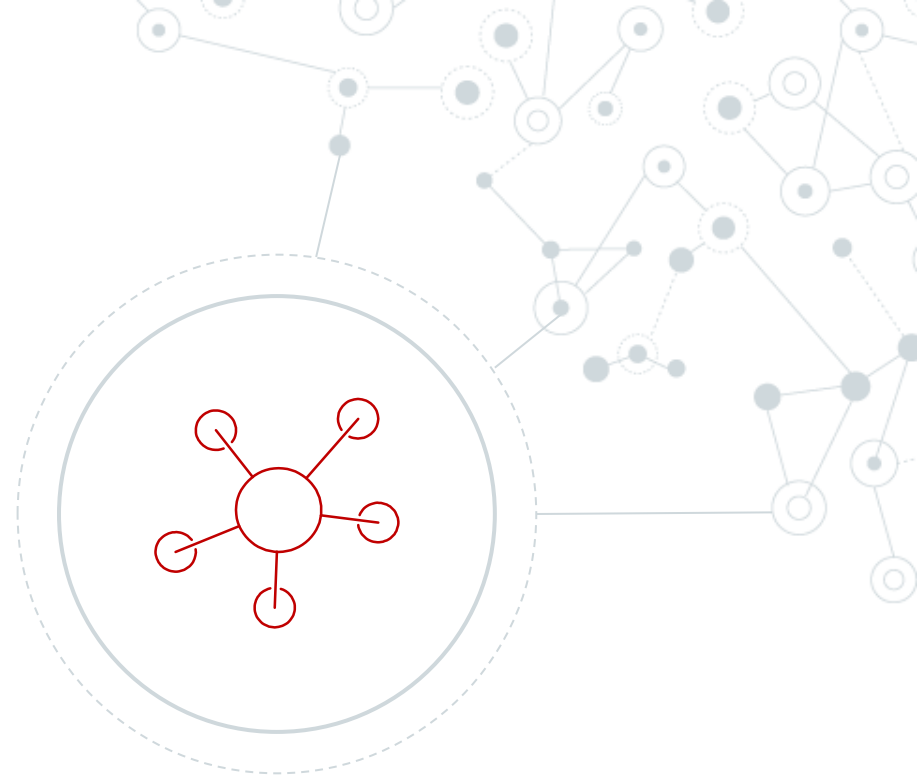
www.iqda.ie



DRI

Digital Repository of
Ireland

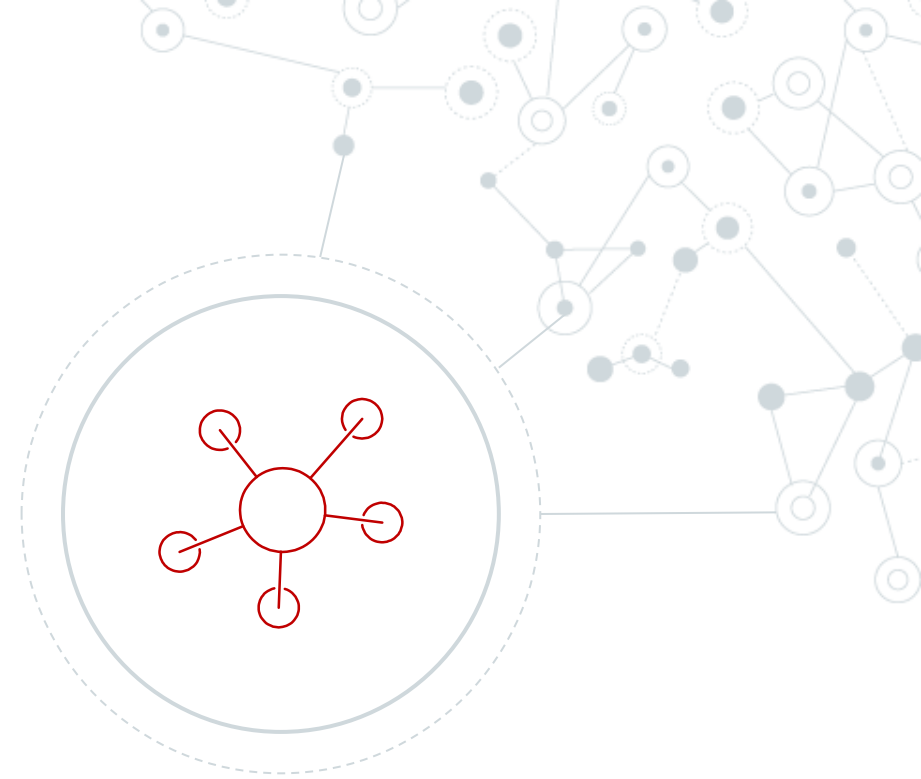
www.dri.ie



CSO

Central Statistics Office
Statbank

<http://www.cso.ie/px/pxeirestat/statire/SelectTable/Omrade0.asp?Planguage=0>



data.gov.ie

Ireland's Open Data
Portal

<https://data.gov.ie/data>



Additional Irish Sources

© Ageing in Ireland www.agestats.ie

© Public health www.thehealthwell.info

© Health www.hiqa.ie/healthcare/health-information/data-collections/online-catalogue

© Children www.dcy.gov.ie/inventory-of-data-sources-on-childrens-lives

Other Archives & Data Sources

Europe

European Social Survey

Social scientific endeavour to map the attitudes, beliefs and behaviour patterns of the various populations in Europe.

www.europeansocialsurvey.org



CSES

Comparative Study of
Electoral Systems

www.cses.org



eurostat

ec.europa.eu/eurostat



STATISTICS BY THEME Statistics A to Z

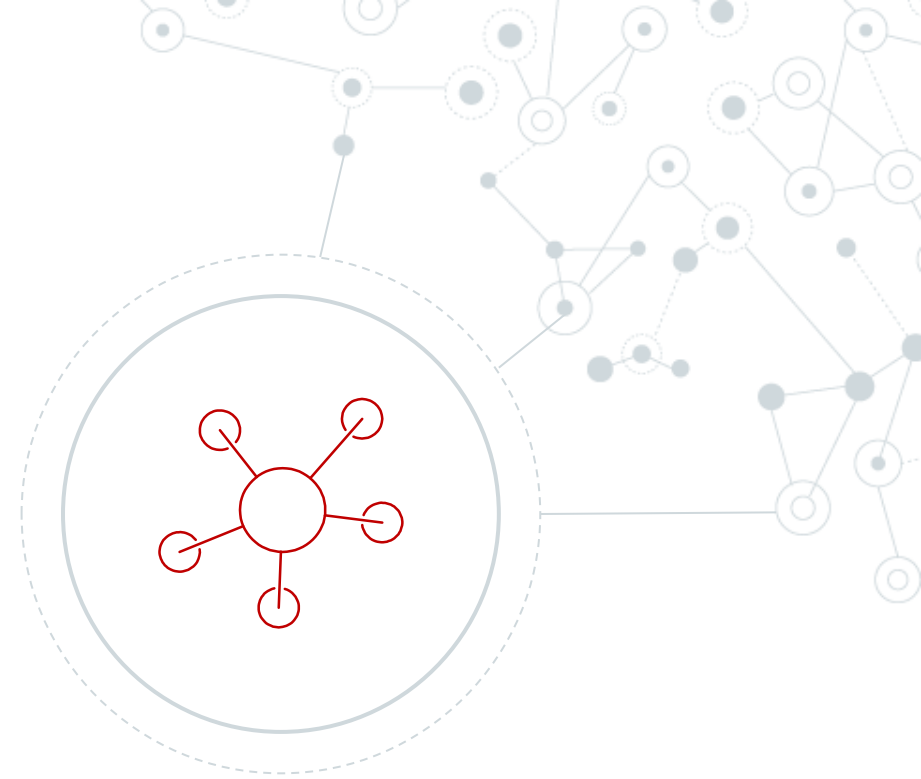
 General and regional statistics	 Economy and finance	 Population and social conditions
 Industry, trade and services	 Agriculture and fisheries	 International trade
 Transport	 Environment and energy	 Science, technology and digital society



UK Data Archive

The UK's largest
collection of digital
research data in the
social sciences and
humanities

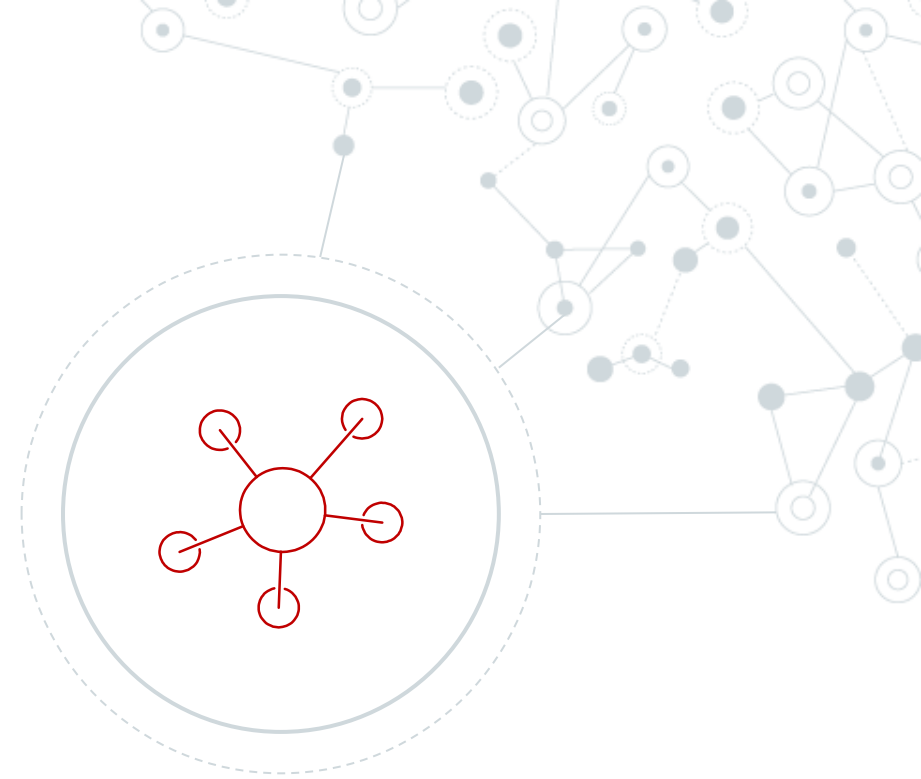
www.data-archive.ac.uk



CESSDA

Council of European
Social Science Data
Archives

www.cessda.org



Other Archives & Data Sources

International

ICPSR

Inter-university
Consortium for
Political and Social
Research

www.icpsr.umich.edu



IPUMS

Integrated Public Use
Microdata Series,
International

<https://international.ipums.org/international/>



International Data Providers

© WHO: <http://www.who.int/gho/database/en/>

© World Bank: <http://data.worldbank.org/>

© <http://www.ucd.ie/issda/help/otherdataarchives/>

4.

Accessing data

Data access: common arrangements

- ◎ Openly available without registration
 - remember to acknowledge source

- ◎ Register (e-mail, university log-in or request user name and password)
 - sometimes register use of data

- ◎ Terms and conditions
 - no redistribution
 - not identifying individuals
 - non-commercial use only or for research and teaching only

Data access: common arrangements

⦿ Order data

- sometimes need data owner permission

⦿ Additional applications for sensitive/confidential

- may only get access in a safe space (dedicated room or remote systems)

⦿ Free

- sometimes charges for commercial use and supplementary services (such as data on CDs)

Accessing Datasets

Identify the dataset(s)
you require

Complete the request
form for research or
teaching purposes

Send
the signed consent
form by email, or to:
ISSDA

http://www.ucd.ie/t4cms/ISSDA_Application_Research_V2.3.docx



NESSTAR



DESCRIPTION TABULATION ANALYSIS

- [-] Restricted Access Studies
 - [-] The All Ireland Traveller Health Survey
 - [-] All Ireland Traveller Health Study
 - [+] Attitudes to Medical Research
 - [+] Attitudes to Poverty
 - [+] Crisis Pregnancy Agency Surveys
 - [-] Diversity in Schools
 - [-] Adapting to Diversity: Irish Schools and Newcomer Students - Primary Level
 - [+] Metadata
 - [-] Variable Description
 - [-] ...
 - id
 - School size grouped
 - Gender mix of school
 - 'Number of Teachers (full-time equivalents)'
 - Have home school liaison officer
 - Have language support teacher
 - Have learning support teacher
 - Have psychologist
 - Have special needs assistant
 - past 5yrs-nos of infant class students
 - other local schools studs may go to
 - all studs accepted
 - criterion-lives in local area
 - criterion-siblings in school
 - criterion-attended early start
 - criterion-date of application
 - criterion-religion
 - criterion-other
 - literacy problems
 - numeracy problems

Dataset: Adapting to Diversity: Irish Schools and
Variable q1_gendermix: Gender mix of

Values	Categories	N	
0	single-sex	97	<div style="width: 13.0%;"></div> 13.0%
1	coed	649	<div style="width: 87.0%;"></div> 87.0%
Sysmiss		286	

SUMMARY STATISTICS

Valid cases	746
Missing cases	286
Minimum	0.0
Maximum	1.0
This variable is numeric	

Cite the data

- © Credit the data creators
- © Allow others to find the data

Northside Partnership; Orla Doyle; UCD Geary Institute PFL Evaluation Team (2017).

Preparing for Life collection: Evaluation of the Preparing for Life early childhood intervention, 2008 - 2015. [collection]. Version 1. Dublin: Irish Social Science Data Archive [distributor] SN: 0055-00. ucd.ie/issda/pfl

Questions

