UCLA Library Data Archive

Doing Data since 1961

Fall Quarter 2017

Today's Topics

- UCLA Library Data Archives: who we are and what we do
 How to search for and access data for research
 Resources, tools and support available to students at UCLA
 Publishing and Data Management Plans

UCLA Library Data Archive

Social Science Data Archive

1120 Rolfe Hall



https://www.library.ucla.edu/location/social-science-data-archive

3/35

Data Archives: Our Services and the Data Life Cycle

The Data Archive provides data services that are tied to the data life-cycle.

Data archiving is a process, not an end state where data is simply turned over to a repository at the conclusion of a study state that the data is strong to see the study. But and study and so so er the course of a project's life cycle and introduce a schedule for dapositing products over the course of a project's life cycle and for the creation and preservation of accurate metadata, ensuring the usubility of the research data itself. Such practices would incorporate archiving as part of the research method." Preserving Research

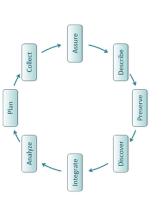


illustration: DatOne Data Life Cycle

Data Archives: Our Services and the Data Life Cycle

https://www.library.ucla.edu/location/social-science-data-archive/data-services

Plan

- Review data sources
 Help investigate archiving issues, costs, consent and disclosure risks
 create a data management plan

Collect/Assure

- Project management and file organization
 Quality Assurance for collected data and acquisition workflows
 Consultation on restricted data access and security

Describe

- Assist with filenames, standard terminology and data dictionaries
 Document analysis and file manipulations
 Help identify appropriate standards

5/35

Data Archives: Our Services and the Data Life Cycle

Preserve

- Help with backing up your Data
 Help with deciding what data to preserve
 Choosing stable file formats
 Help identifying suitable repositories

Integration

- Consider compatability
 Document steps
 Capture the provenance of sources

Analysis

- Help identify appropriate sofware
 Consultation on coding best practices

Data Archives: Our Services and Software

• Provide training: Software/Library carpentry / Software Carpentry R-





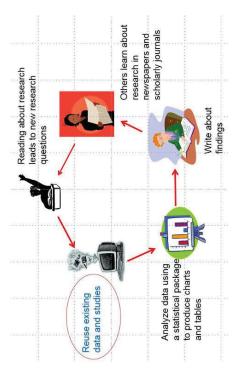




- Host events and lectures: Data Archive News and Upcoming Events
- Statistical Consulting: Refer to UCLA Statistics Consulting Center

7 / 35

Role of Data in Research: Data Reuse



Finding Out About Data Sources

- UCLA Library Data Archive
 Colleagues
 Public media (newspapers, news)
 Social media (twitter, etc)
 Scholarly literature
 Government reports and documents
 Archives and project websites
 Codebooks, questionnaires
 Exploration of datasets

Data Sources

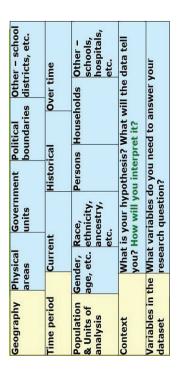
- ICPSR: http://www.icpsr.umich.edu/icpsrweb/ICPSR/

 UCLA Library Data Archive Catalog:
 https://dataverse.harvard.edu/dataverse/ssda_ucla

 UCLA Library DataDen: https://dataden.library.ucla.edu/jspui/
 Government Sites, Census Data: https://www.census.gov/
 IPUMS (Integrated Public Use Microdata Series): https://www.ipums.org/
 Public Policy Institute of CA: http://www.ppic.org/
 NBER (National Bureau of Economic Research): http://www.nber.org/data/
 CDC/NCHS (Centers for Disease Control and Prevention/National Center for Health Statistics): https://www.cdc.gov/nchs/surveys.htm
 Pew Research Center: http://www.pewresearch.org/
 Data.gov: https://www.data.gov/
 Norld Health Organization: http://www.inf.org/en/Data

Search Strategies and Defining Your Research

- General to specific; broad to narrow
 Names of investigators; names of studies
 Studies described in literature, news and websites



11/35

How do I know this study useful?

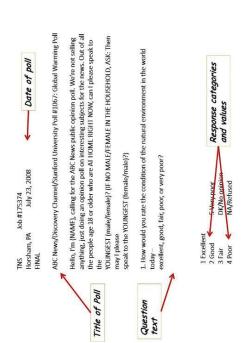
- Raw data are not eye-readable To decipher data you will use ${\bf Question naires}$ and Codebooks

Dump of Data --This is an example of a row duta file of a compus computing nerwook

What is a Questionnaire?

- A questionnaire is a research instrument consisting of a series of questions and other prompts for the purpose of gathering information from respondents.
- Questionnaires cover separate topics such as:
- Preferences (e.g. political party)
 Behaviors (e.g. food consumption)
 Facts (e.g. gender)
- Questionnaires can include indexes or scales:
- Latent traits (e.g. personality traits)
 Attitudes (e.g. towards immigration)
 An index (e.g. Social Economic Status)

Questionnaire: ABC New Poll



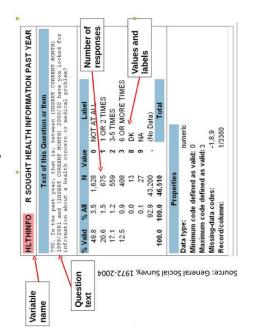
What is a Codebook?

- An overview of the study or dataset
 Variable names and descriptions
 Column locations (mostly this is for raw data)

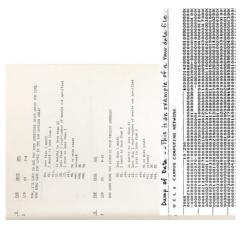
May also include:

- Test of questions if the data comes from a survey
 Example of the Questionnaire
 Detail on who responded to the survey

Codebook: General Social Survey



Codebook and Raw Data



17/35

ICPSR Data Site

Inter-university Consortium for Political and Social Research $\ensuremath{\mathsf{http://www.icpsr.umich.edu}}$

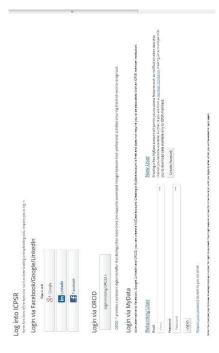


ICPSR: Find Data



ICPSR: Register to Access Data

(ORCID Identifier: https://orcid.org/)



ICPSR: Search by Topics

Browse by Topic

Topic Considerate alemping and

Topic Classification and in the median in the media

21 / 35

ICPSR: Search for and Compare Variables

Search/Compare Variables
The sucidification that the sucidification of the sucidificatio

Comparing Variables

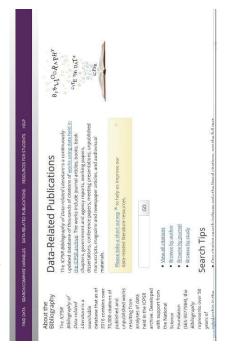
ICPSR: Compare Variables



ICPSR: Study Page and Data

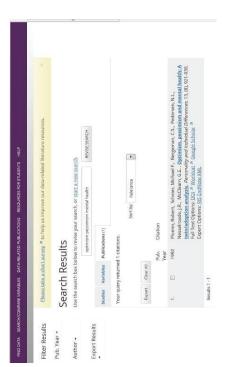


ICPSR: Reading About Data



25/35

ICPSR: Publications linked to Data



ICPSR: Publications Linked to Data

Please take a short survey A to help us improve our data-related literature resources.

Citation

The following citation is from the ICPSR Bibliography of Data-related Literature:

Plomin, Robert, Scheler, Michael E., Bergeman, C.S., Pedersen, N.L., Nesselroade, J.R., McCleam, G.E., Aug 1992. Opt pessimism and mental health: A twin/adoption analysis. *Personality and Individual Differences*, 13, (8), 921-930. doi:10.1016/0191-886992)90009-E. Full Text Options: DOI ³ Worldcat. ³ Google Scholar. ³ Export Citation: BIS EndNote XML

Related Studies

This publication is related to the following dataset(s):

• Swedish Adoption/Twin Study on Aging (SATSA), 1984, 1987, 1990, 1993, 2004, 2007, and 2010

27/35

Publishing and Data Management Plans

- What is a Data Management Plans
 Why do I need this?
 What does this have to do with Publishing?
 What tools are available to help?

What is a Data Management Plan?

A data management plan is a document that describes what you will do with your data during your research and after you complete your research.

29 / 35

Why do you need this?

- Funding Agencies
 Publishing
 Sharing and Archiving Data
 Regardless of the requirements, good data management is an essential skill for researchers.

Example: Williams Institute, UCLA School of Law / UCLA DataDen

Basic Elements of a Data Management Plan

- What are your data?
 What formats will you be using?
 How will you describe this data?
 What intellectual property and privacy rights are associated with this data? (Will you need IRB approval?)
 How will you share this data? If you don't plan on sharing it, why not?
 How much will your data management cost?

See also ICPSR Elements of a Data Management Plan

Getting Help with Data Management Plans

ICPSR Resources for Data Management

- ICPSR Framework for Creating a Data Management Plan
 Data Management and Curation
 ICPSR Guide to Archiving Social Science Data for Institutional Repositories, 5th edition

Other Resources

- UCLA Data Archive
 UCLA Library: Research Guides
 UKDA tools for creating and managing data (Pdf document)
 MANTRA online learning tool/tutorial
 JISC Digital Media Advice

Data Management and Confidentiality

If your research involves human subjects, you will need to consider both legal and ethical obligations in managing and sharing your data.

Confidentiality refers to the agreement between the researcher and the participant about how the participant's identifiable private information will be handled, managed, and disseminated.

As a researcher, you need a clear view about how to protect the privacy of your research subjects.

33 / 35

Questions?



Problems?

The Data Archive is here to help

1120-H Rolfe ucla-data-archive@googlegroups.com 310-825-0716 Tim Dennis, Director of the Data Archive Jamie Jamison, Archive Staff

presentation