

Using Data in Research

UCLA Library Data Archive

Fall 2017

Today's Topics

- ▶ UCLA Library Data Archives
- ▶ The role of data in your research
- ▶ How to search for and access data for research
- ▶ Resources, tools and support available to students at UCLA
- ▶ Publishing and Data Management Plans

Role of Data in Research

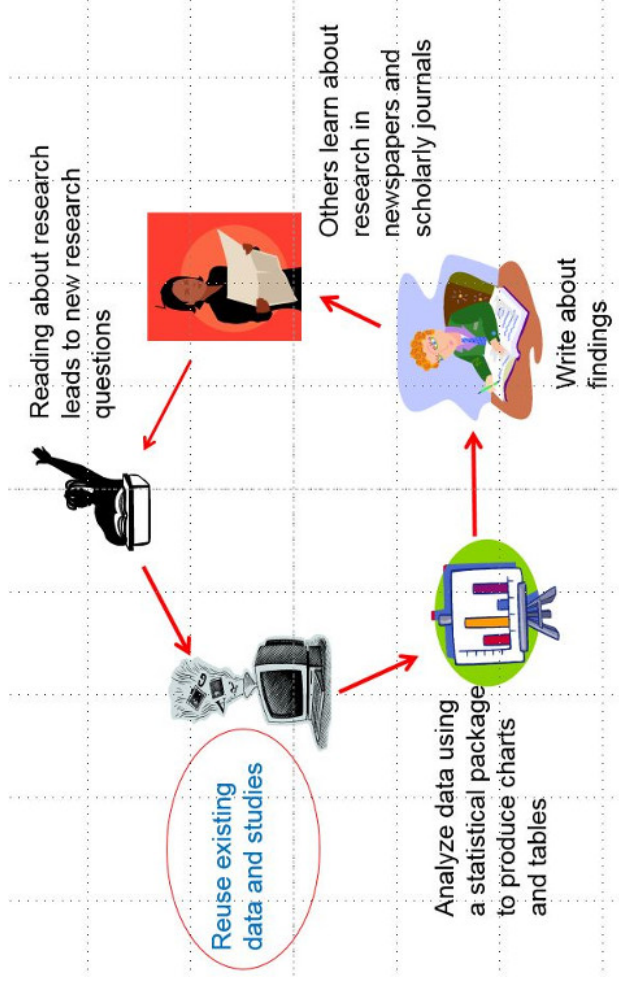


Figure 1: illustration: Role of Data in Research

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

Search Strategies

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

Geography	Physical areas	Government units	Political boundaries	Other – school districts, etc.	
Time period	Current		Historical	Over time	
	Population & Units of analysis	Gender, age, etc.	Race, ethnicity, ancestry, etc.	Persons	Households
Context	What is your hypothesis? What will the data tell you? How will you interpret it?				
Variables in the dataset	What variables do you need to answer your research question?				

How do I know this study useful?

- ▶ Raw data are **not** eye-readable
- ▶ To decipher data you will use **Questionnaires and Codebooks**

Dump of Data -- This is an example of a raw data file.

U C L A C A M P U S C O M P U T I N G : N E T W O R K

[illegible]

Figure 2: illustration: Raw Data

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

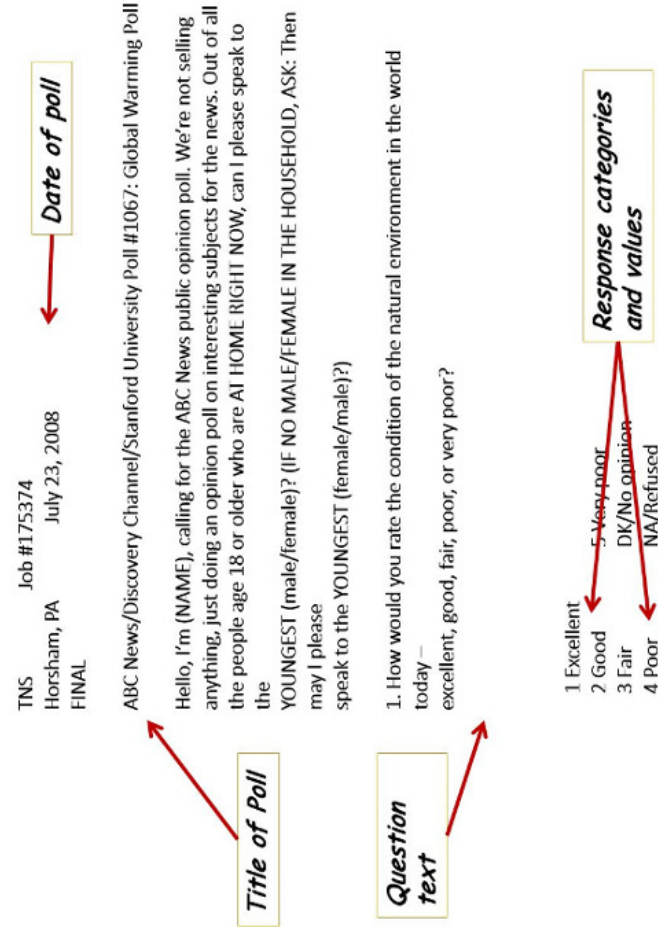
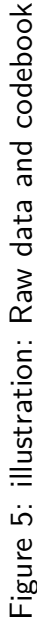


Figure 3: illustration: ABC News Poll

H	LTHINFO	R SOUGHT HEALTH INFORMATION PAST YEAR			
Text of this Question or Item					
792. In the past year, that is, between (INSERT CURRENT MONTH) 1999/2001 and (INSERT CURRENT MONTH) 2000/02 have you looked for information about a health concern or medical problem?					
% Valid	% All	N	Value	Label	Number of responses
49.8	3.5	1,628	0	NOT AT ALL	
20.6	1.5	675	1	1 OR 2 TIMES	
17.1	1.2	559	2	3-5 TIMES	
12.5	0.9	408	3	6 OR MORE TIMES	
	0.0	13	8	DK	
	0.1	27	9	NA	
	92.9	43,200	.	(No Data)	
100.0	100.0	46,510		Total	
Data type:					
Minimum code defined as valid: 0					
Maximum code defined as valid: 3					
Missing-data codes: -1,-8,9					
Record/column: 1/3360					

Source: General Social Survey, 1972-2004

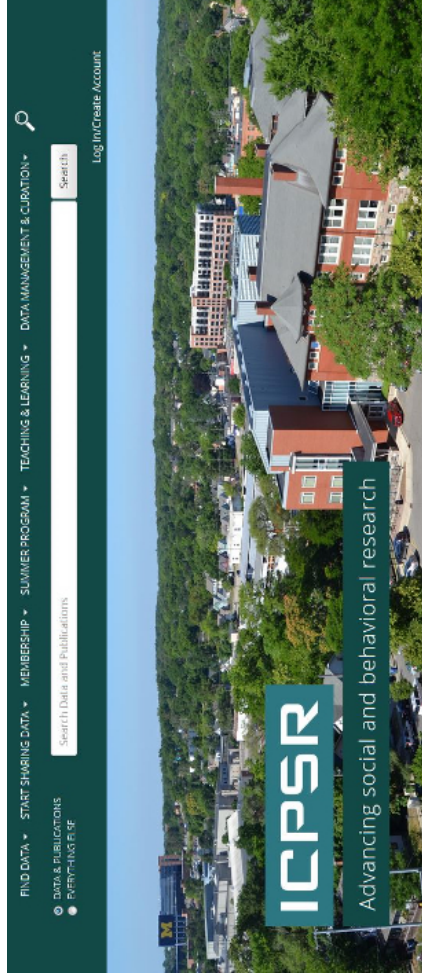
Codebooks: Example #2



ICPSR Data Site

Inter-university Consortium for Political and Social Research

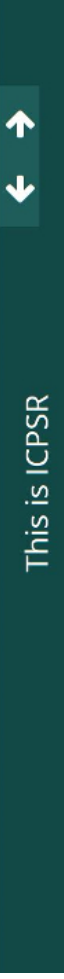
<http://www.icpsr.umich.edu>



News

More Headlines

[view all news](#)



ICPSR: Find Data

<http://www.icpsr.umich.edu/icpsrweb/ICPSR>



Find Data

More Search Options

- Browse by topic
- Browse by series
- Browse by geography
- Browse by investigator
- Browse by data collection
- Browse international data
- View all studies
 - View all studies for which online analysis is available
 - View all replication datasets (online only)
 - View all studies that have learning guides

Restricted-Use Data

ICPSR ensures respondent confidentiality within public use versions of the datasets. The vast majority of ICPSR data holdings are public-use files with no access restrictions. Sometimes the protective measures taken to reduce disclosure risk would significantly degrade the research potential of the data. In these cases, ICPSR provides access to restricted use versions that retain confidential data by imposing stringent requirements for accessing them. (U.S.)

Search Tips

Select a link below to reveal search tips relevant to your search.

[I'm looking for datasets related to a specific topic.](#)

[I'm looking for datasets that contain specific variables.](#)

[I'm looking for datasets that answer a particular research question.](#)

[I'm looking for a specific dataset. I know the title and/or principal investigator.](#)




ICPSR: Register to Access Data

Log into ICPSR


Some functions of the web site, such as downloading and uploading data, require you to log in.

Login via Facebook/Google/LinkedIn


Sign in with



Google



LinkedIn



Facebook

Login via ORCID

Sign in Using ORCID >

ORCID >

provides a persistent digital identifier that distinguishes researchers and supports automated linkages between their professional activities ensuring that their work is recognized.

Login via MyData

As an alternative to Facebook, Google, LinkedIn and ORCID, you can create a MyData account. Creating a MyData account is free and does not require you to be associated with an ICPSR member institution.

New User

Creating a free MyData account will permit you to access features such as notification when data that interest you become available. Further, all data are from a [member institution](#), creating an account permits you to download data available only to ICPSR members.

Email:

Password:

Log In

Create Account

Returning User

Email:

Password:

Log In

Requires your password be sent to you via email.

After you make our password be e-mailed to you to log in to account. Your login name will last for one hour (or until you log out), after which you will be asked to log in again.

Figure 7: illustration: ICPSR Register

ICPSR: Search by Topics

Browse by Topic

Users have two options for identifying studies related to a particular topic.

- [Topic Classifications](#) address topics broadly, and are meant to be a good starting point for exploring the collection.
- The [Thesaurus](#) contains both broad and narrow terms that are more specific to the subject matter of individual studies.
- Our [Thematic Collections](#) are full websites that focus on particular subject areas.

Topic Classifications

- I. Census Enumerations: Historical and Contemporary Population Characteristics
 - A. United States
 - 1. Decennial Censuses
 - a. 1790-1960 Censuses
 - b. 1970 Census
 - c. 1980 Census
 - d. 1990 Census
 - e. 2000 Census
 - f. 2010 Census
 - 2. American Housing Survey Series
 - 3. Current Population Survey Series
 - 4. Other: Census, Including County and City Data Books
 - B. Nations Other Than the United States
- II. Community and Urban Studies
 - A. Studies of Local Politics
 - 1. United States
 - 2. Nations Other Than the United States
 - B. Detroit Area Studies
 - C. Historical Urban and Community Structure Studies
- III. Conflict, Aggression, Violence, Wars
 - A. Conflict and Stability Within Nations
 - B. Conflicts Between and Among Nations
- IV. Economic Behavior and Attitudes
 - A. [Productivity Growth and Performance: A Review](#)

Figure 8: illustration: ICPSR Search by Topics

ICPSR: Search for and Compare Variables

Search/Compare Variables

The Social Science Variables Database (SSVD) enables ICPSR users to examine and compare variables and questions across studies or series. The SSVD currently includes over 4 million variables, representing about 76% of ICPSR's holdings that have quantitative data described in statistical syntax.

The SSVD does not include our entire collection. For a comprehensive search of all datasets, use our [main data search](#), which indexes all of the documentation available for our studies, including the variable descriptions.

college major	GO
---------------	----

- [Browse by Study](#)
- [Browse by Series](#)

Comparing Variables

Compare variables using the checkboxes and the "Compare" button on the search results page.

Results per page: 25	Results per page: 25
<p>Results</p> <p>1</p> <p>2</p> <p>3</p> <p>4</p> <p>5</p> <p>6</p> <p>7</p> <p>8</p> <p>9</p> <p>10</p> <p>11</p> <p>12</p> <p>13</p> <p>14</p> <p>15</p> <p>16</p> <p>17</p> <p>18</p> <p>19</p> <p>20</p> <p>21</p> <p>22</p> <p>23</p> <p>24</p> <p>25</p>	<p>Results</p> <p>1</p> <p>2</p> <p>3</p> <p>4</p> <p>5</p> <p>6</p> <p>7</p> <p>8</p> <p>9</p> <p>10</p> <p>11</p> <p>12</p> <p>13</p> <p>14</p> <p>15</p> <p>16</p> <p>17</p> <p>18</p> <p>19</p> <p>20</p> <p>21</p> <p>22</p> <p>23</p> <p>24</p> <p>25</p>

Search Tips

- Enter words or strings that are likely to appear in a variable name, label, question, and value labels:
- Presidential election* will return variables dealing with all presidential elections
- Presidential Election Outcome* will return only variables dealing with the 2008 and 2012 presidential elections
- Use quotes to search for specific phrases:
- "life satisfaction"
"minority rights"
"community programs"
- The minus sign may be used to remove certain types of results:
- "Presidential debate" - "Bill Clinton" will eliminate the debates from 1992 and 1996 in which Clinton participated
- A Boolean "and" is implied in the search.
 - The search automatically does stemming ([link](#) [info](#)). If there's no need to turn an asterisk, it's also

Figure 9: illustration: ICPSR Search for and Compare Variables

ICPSR: Compare Variables

Compare Variables

Compare (3) Clear All (3) Back to search results

<input checked="" type="checkbox"/> MAJOR1	<input checked="" type="checkbox"/> MAJOR2	<input checked="" type="checkbox"/> OC12												
General Social Survey, 1972-2014 (Cumulative File)	General Social Survey, 2012 Merged Data, Including a Cultural Module (United States)	Longitudinal Study of American Youth, 1987-1994, 2007-2011: LSAY Merged Cohort File (Base File)												
COLLEGE MAJOR 1	COLLEGE MAJOR 2	COLLEGE MAJOR												
Question	Question	Question												
"What was your major or field of study when you received your (RESPONDENT'S COLLEGE DEGREE) degree? IF THE RESPONDENT RECEIVED MORE THAN ONE GRADUATE LEVEL DEGREE, ASK ABOUT THE HIGHEST DEGREE OBTAINED/ COUNT UP TO TWO MENTIONS."	What was your major or field of study (RESPONDENT'S COLLEGE DEGREE) degree? Added text: IF THE RESPONDENT RECEIVED DEGREE, ASK ABOUT THE HIGHEST DEGREE RATHER THAN AN M.B.A.; [COUNT UP TO 2													
Responses	Responses	Responses												
<table><tr><th>VALUE</th><th>LABEL</th></tr><tr><td>0</td><td>Unapplicable</td></tr></table>	VALUE	LABEL	0	Unapplicable	<table><tr><th>VALUE</th><th>LABEL</th></tr><tr><td>0</td><td>Inapplicable</td></tr></table>	VALUE	LABEL	0	Inapplicable	<table><tr><th>VALUE</th><th>LABEL</th></tr><tr><td>-99</td><td>Not asked</td></tr></table>	VALUE	LABEL	-99	Not asked
VALUE	LABEL													
0	Unapplicable													
VALUE	LABEL													
0	Inapplicable													
VALUE	LABEL													
-99	Not asked													

Figure 10: illustration: ICPSR Compare Variables

ICPSR: Study Page and Data

THE ICPSR: SEARCHING FOR DATA-RELATED PUBLICATIONS: RESOURCES FOR STUDENTS HELP

High School Longitudinal Study, 2009-2013 [United States] (ICPSR 36423)

Funded by...

Alternate Title: HSLS09

Principal Investigator(s) for the Study: Department of Education, Institute of Education Sciences, National Center for Education Statistics

Summary

The High School Longitudinal Study of 2009 (HSLS09) is a national, representative, longitudinal study of 9th graders in the United States. The study was designed to provide information on the educational experiences and outcomes of students in the United States. The study was designed to provide information on the educational experiences and outcomes of students in the United States.

Table of Contents

- About HSLS
- About HSLS09
- About HSLS09 Data
- About HSLS09 Data Files
- About HSLS09 Data Files
- About HSLS09 Data Files

Access Notes

- This study is available in the public domain.

Datasets

- HSLS09 Data Files
- HSLS09 Data Files
- HSLS09 Data Files

ICPSR Study-Level Files

Documentation: Documentation of the Study

Download: Download the Study

Analyze: Analyze the Study

Online: Online the Study

Table of Contents

- About HSLS
- About HSLS09
- About HSLS09 Data
- About HSLS09 Data Files
- About HSLS09 Data Files
- About HSLS09 Data Files

Study Description

Citation

Figure 11: illustration: ICPSR Study Page and Data

ICPSR: Reading About Data

FIND DATA SEARCH/COMPARE VARIABLES DATA-RELATED PUBLICATIONS RESOURCES FOR STUDENTS HELP

About the Bibliography

The ICPSR Bibliography of Data-related Literature is a continuously updated database of thousands of citations of works using data held in the ICPSR archive. The works include journal articles, books, book chapters, government and agency reports, working papers, dissertations, conference papers, meeting presentations, unpublished manuscripts, magazine and newspaper articles, and audiovisual materials.

Data-Related Publications

Please take a [short survey](#) to help us improve our data-related literature resources.

GO

Search Tips

- [View all citations](#)
- [Browse by author](#)
- [Browse by journal](#)
- [Browse by study](#)

Our citation search includes only the literal citations, not the full text.

Figure 12: illustration: ICPSR Reading About Data

ICPSR: Publications Linked to Data

FIND DATA | SEARCH/COMPARE VARIABLES | DATA-RELATED PUBLICATIONS | RESOURCES FOR STUDENTS | HELP

Please take a [short survey](#) to help us improve our data-related literature resources.

Filter Results

Pub. Year ▾

Author ▾

Export Results ▾

Search Results

Use the search box below to revise your search, or [start a new search](#).

optimism pessimism mental health

REVISE SEARCH

Studies

Variables

Publications (1)

Your query returned 1 citations.

Sort by: Relevance ▾

Export Clear All

1.

1992

Plomin, Robert, Scheier, Michael F., Bergeman, C.S., Pedersen, N.L., Nesselroade, J.R., McClearn, G.E. **Optimism, pessimism and mental health: A twin/adoption analysis.** *Personality and Individual Differences* 13, (8), 921-930. [Full Text Options: DOI](#) [WorldCat](#) [Google Scholar](#) [Export Options: RIS](#) [EndNote XML](#)

Results 1 - 1

Figure 13: illustration: ICPSR Publications linked to Data

ICPSR: Publications Linked to Data

FIND DATA | SEARCH/COMPARE VARIABLES | DATA-RELATED PUBLICATIONS | RESOURCES FOR STUDENTS | HELP

Please take a [short survey](#) to help us improve our data-related literature resources.

Citation

The following citation is from the [ICPSR Bibliography of Data-related Literature](#):

Plomin, Robert, Scheier, Michael F., Bergeman, C.S., Pedersen, N.L., Nesselroade, J.R., McClearn, G.E. .Aug 1992. Optimism, pessimism and mental health: A twin/adoption analysis. *Personality and Individual Differences* 13, (8), 921-930. doi:10.1016/0191-8869(92)90009-E

Full Text Options: [DOI](#) [WorldCat](#) [Google Scholar](#)

Export Citation: [RIS](#) [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](#)

Figure 14: illustration: ICPSR Publications Linked to Data

Data Sources

- ▶ ICPSR: <http://www.icpsr.umich.edu/icpsrweb/ICPSR/>
- ▶ UCLA Library Data Archive Catalog: https://dataverse.harvard.edu/dataverse/ssda_ucla
- ▶ Government Sites, Census Data: <https://www.census.gov/>
- ▶ IPUMS (Integrated Public Use Microdata Series): <https://www.ipums.org/>
- ▶ Public Policy Institute of CA: 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/>

Data Management Plans

- ▶ What is a Data Management Plans
- ▶ Why do I need this?
- ▶ 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.

Why do you need this?

- ▶ Funding Agencies
- ▶ Publishing <https://datadn.library.ucla.edu/jspui/>
ORCID:<https://orcid.org/>
- ▶ Sharing and Archiving Data
- ▶ Regardless of the requirements, good data management is an essential skill for researchers.

Getting help

- ▶ Data Archive
- ▶ ICPSR

Questions



Figure 15: illustration: 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