Using Data in Research

Library Data Archive Fall Quarter 2017

PP290 Public Policy Fall 2017

1/37

Today's Topics

- UCLA Library Social Science 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

UCLA Library Data Archives

- On the UCLA campus since 1961 About the Data Archive
 Available to all UCLA faculty and students
- Holdings include public opinion polls, census data, crime stats, political and social behavior, etc.



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

3/37

Data Archive: Our services

- Reference: assist students to find and use data
- Data Management Plans
- Metadata
- Provide training: Software/Library carpentry / Software Carpentry R-language

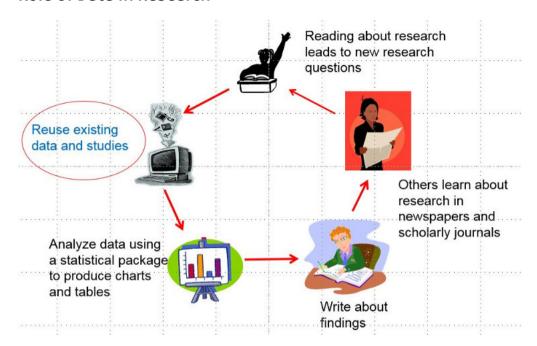






• Host events and lectures: Data Archive News and Upcoming Events

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 documentsArchives and project websites
- Codebooks, questionnaires
- Exploration of datasets

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

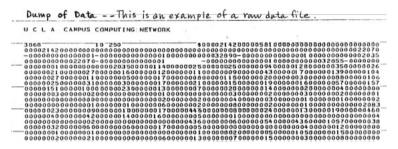
Geography	Physical areas		Governr units	Government units		Political boundaries		Other – school districts, etc.	
Time period	e period C		Current		Historical		Over time		me
	age, etc.		Race, ethnicity, ancestry, etc.	Persons		Household		ds Other – schools, hospitals, etc.	
Context		What is your hypothesis? What will the data tell you? How will you interpret it?							
Variables in dataset		What variables do you need to answer your research question?							

Questionnaires and Codebooks

How do I know this study useful?

Studies with Raw Data

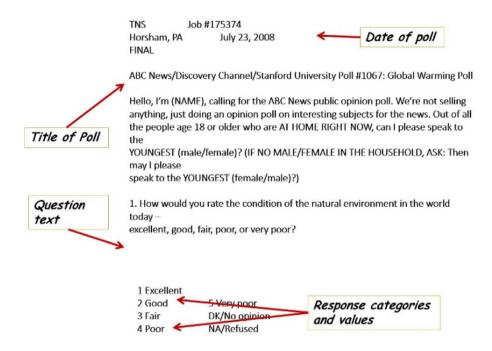
- Raw data are **not** eye-readable
- To decipher data you will use Questionnaires and Codebooks



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



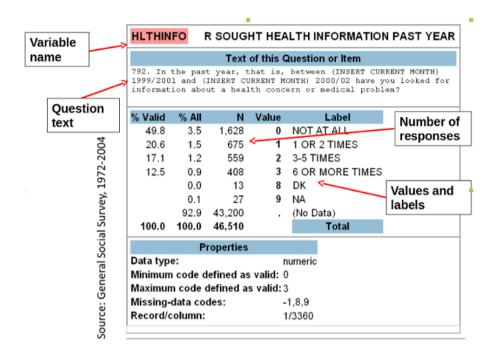
What is a Codebook?

- 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

```
9
         VAR
                   DECK
          119
                   05
         NOW, I'D LIKE TO ASK YOU SOME QUESTIONS ABOUT WHERE YOU LIVE. HOW LONG HAVE YOU LIVED IN THE LOS ANGELES AREA?
                     00. less than 1 month
01. 1 month to less than 2
                    etc.
11. 11 months to less than 12
                     12. less than 1 year-number of months not specified
13. 1 year to less than 2
                    96. 96 or more years
*97. refused
*98. IK
                    498.
                    499. NA
         VAR
                   DECK
                             COL
                   05
                            9-10
         HOW LONG HAVE YOU LIVED AT YOUR PRESENT ADDRESS?
                     00. less than 1 month
01. 1 month to less than 2
                     11. 11 months to less than 12
12. less than 1 year-number of months not specified
11. 1 year to less than 2
     Dump of Data -- This is an example of a raw data file
    U C L A CAMPUS COMPUTING: NETWORK
```

ICPSR Data Site

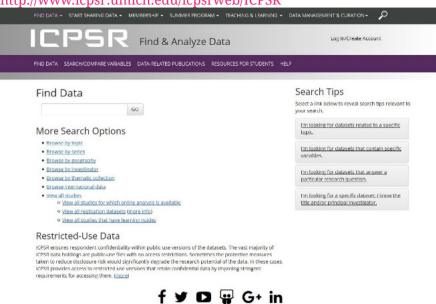
Inter-university Consortium for Political and Social Research

http://www.icpsr.umich.edu

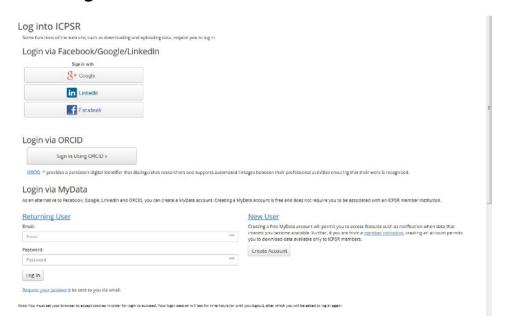


ICPSR: Find Data

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



ICPSR: Register to Access Data



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

IV. Economic Behavior and Attitudes

```
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. Conflict Between and Among Nations
```

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 source.

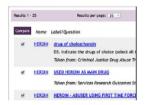
The SSVD does not include our entire collection. For a comprehensive search of all datasets, use our <u>main data search</u>, which indexes all of the documentation available for our studies, including the variable descriptions.



- · Browse by Study
- Browse by Series

Comparing Variables

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



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 Obama 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"

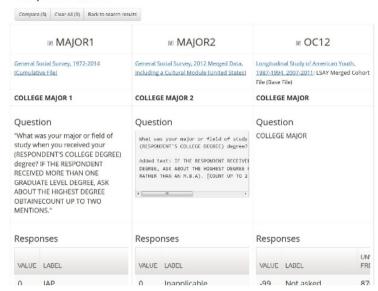
 The minus sign may be used to remove certain types of results:

"Presidential debate" - "Bill Clinton" will eliminate the debates from 1992 and 1996 in which Clinton participated

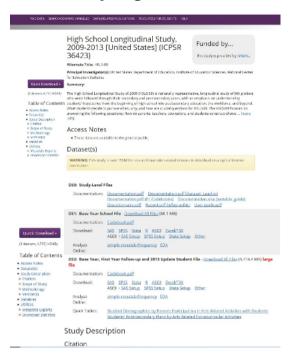
- A Boolean "and" is implied in the search.
- The search automatically does stemming (more info ²⁸); there's no need to type in an asterisk. It's also

ICPSR: Compare Variables

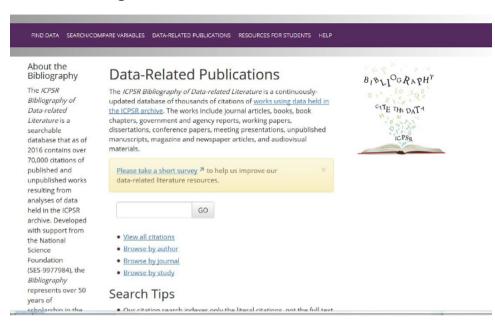
Compare Variables



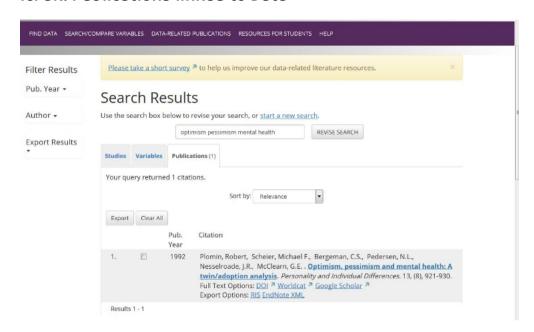
ICPSR: Study Page and Data



ICPSR: Reading About Data



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 7 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. Opt 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

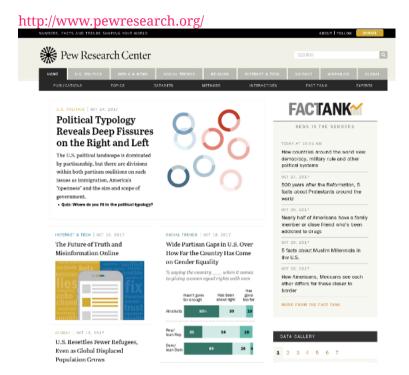
Export Citation: RIS EndNote XML

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

Pew Research Center



Roper Center: Articles, Tables and Access to Data

https://ropercenter.cornell.edu/



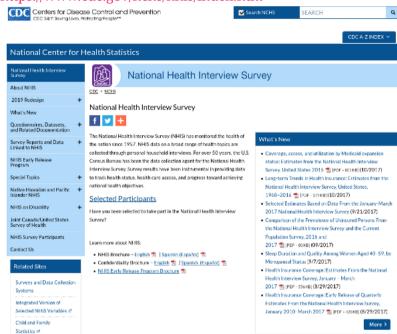
Question Search

The Pew Research Center for the People & the Press Poll Database



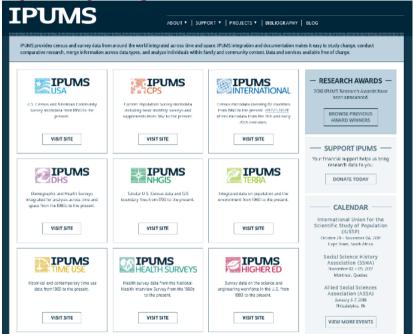
National Center for Health Statistics

https://www.cdc.gov/nchs/nhis/index.htm



IPUMS: Integrated Health Interview Series

https://www.ipums.org/



National Bureau of Economic Research

http://www.nber.org/data/



Data from U.S. Census Bureau

Census

https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml



Data Sources

- ICPSR: http://www.icpsr.umich.edu/icpsrweb/ICPSR/
- UCLA Library Data Archive Catalog: https://dataverse.harvard.edu/dataverse/ssda_ucla
- Roper Center for Public Opinion Research: https://ropercenter.cornell.edu/
- 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 PlansWhy 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://dataden.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 with Data Management Plans

- Data ArchiveICPSR

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