

# Using Data in Research

Library Data Archive Fall Quarter 2017

Psych M144/Stats M154 Fall 2017

1 / 30

## 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

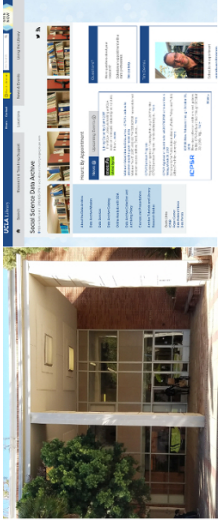
Psych M144/Stats M154 Fall 2017

2 / 30

## UCLA Library Data Archives

(formerly the Social Science Data Archive)

- 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>

Psych M144/Stats M154  
Fall 2017

3 / 30

## 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**

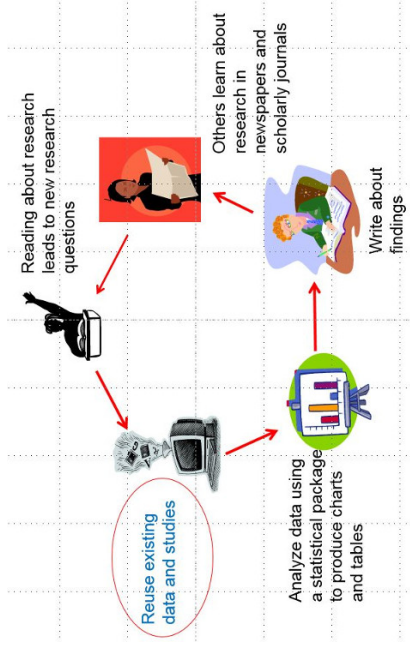


- Host events and lectures: **Data Archive News and Upcoming Events**

Psych M144/Stats  
M154 Fall 2017

4 / 30

## Role of Data in Research



Psych M144/Stats M154 Fall 2017

5 / 30

## 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

Psych M144/Stats M154 Fall 2017

6 / 30



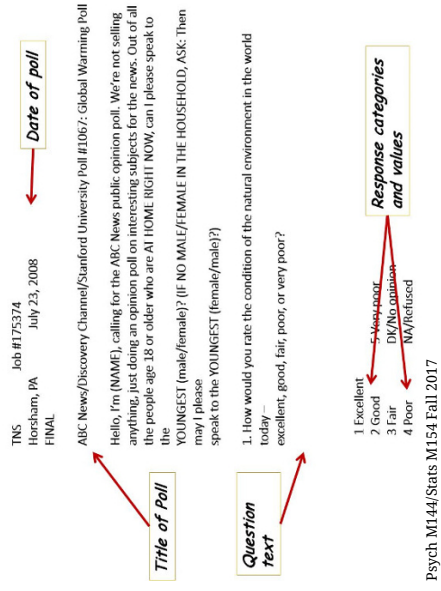
## What is a Questionnaire?

- 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)

Psych M144/Stats M154 Fall 2017

9 / 30

## Questionnaire: ABC New Poll



10 / 30

## 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

Variable name	HLTHINFO	R SOUGHT HEALTH INFORMATION PAST YEAR				
Question text	Text of this Question or Item 1999/2001 and 1999/2001 and 1999/2001 and 1999/2001 have you looked for information about a health concern of medical problem?					
	<b>% Valid</b>	<b>% All</b>	<b>N</b>	<b>Value</b>	<b>Label</b>	<b>Number of responses</b>
	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)	
	<b>100.0</b>	<b>100.0</b>	<b>46,510</b>		<b>Total</b>	
	<b>Properties</b>					
	Data type: numeric					
	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



# ICPSR: Find Data

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

ICPSR Find & Analyze Data

ICPSR: SEARCHING RESULTS CONTACT US ABOUT US

## Find Data

### More Search Options

- Browse by Year
- Browse by Country
- Browse by Discipline
- Browse by Method
- Browse by Year of Data
- View all datasets
  - All of the datasets that are available in our archive
  - All of the datasets that have been archived

### Restricted-Use Data

ICPSR ensures respondent confidentiality within public use versions of the datasets. The vast majority of datasets are available in public use versions. However, some datasets are available only in restricted-use versions. These datasets are available only to researchers who have been approved by ICPSR. For more information, please contact us at [icpsr@umich.edu](mailto:icpsr@umich.edu).



### Search Tips

For more information about searching for relevant to your search:

- To obtain the contact information for a specific dataset, click on the dataset's title.
- To obtain the contact information for a specific dataset, click on the dataset's title.
- To obtain the contact information for a specific dataset, click on the dataset's title.
- To obtain the contact information for a specific dataset, click on the dataset's title.

Psych M144/Stats M154 Fall 2017

# ICPSR: Register to Access Data

## Log into ICPSR

Log in with your account, such as those created with Google, Facebook, or LinkedIn.

Login Via Facebook/Google/LinkedIn

Log in with:

- Facebook
- Google
- LinkedIn
- Facebook

## Login via ORCID

Sign in using ORCID iD.

ORCID is a persistent digital identifier that allows you to connect and explore relationships between your professional activities, ensuring that your work is recognized.

## Login via MyData

Log in with your MyData account. MyData is a secure, cloud-based data management system that allows you to store, manage, and analyze your data.

## Returning User

Enter your email address and password to log in. If you are having trouble logging in, please contact us at [icpsr@umich.edu](mailto:icpsr@umich.edu).

Email:

Password:

Remember your password to auto-login to your account.

Forgot your password? Please contact us at [icpsr@umich.edu](mailto:icpsr@umich.edu) for assistance. Your login session will end when you log out or your session expires.

Psych M144/Stats M154 Fall 2017



# ICPSR: Search by Topics

## Browse by Topic

Users who are interested in finding studies related to a particular topic

- Topic Classification address topics broadly, and are meant to be a good starting point for exploring the collection.
- The Thesauri present both broad and narrow terms that are more specific to the subject matter of individual studies.
- Our Electronic Collections are full subjects that focus on specific questions or subject areas.

## Topic Classifications

- I. General Characteristics: Historical and Contemporary Population Characteristics
  - A. International Countries
    - 1. 1950-1959 Census
    - 2. 1960 Census
    - 3. 1980 Census
    - 4. 1990 Census
    - 5. 2000 Census
  - B. 2000 Census
  - C. 1998 Census
  - D. 1992 Census
  - E. 2002 Census
- 2. American Historical Topics
  - 3. Current Population Survey Series
  - 4. Other Census, Including County and Data Banks
- II. Historical Topics: 1980-2000
  - A. Studies of Local Politics
    - 1. United States
    - 2. International
  - B. Studies of State Politics
    - 3. Historical Urban and Community Structure Studies
  - C. Historical Urban and Community Structure Studies
  - III. Scientific Approaches: Methods, Models, and Techniques
    - A. Quantitative Methods and Analytical Techniques
    - B. Content Analysis and Archival Methods
    - IV. Economic Indicators and Methods
      - 1. Production, Quality of Production, G. Growth

Psych M144/Stats M154 Fall 2017

# 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. 70% of ICPSR holdings that have quantitative data specified in statistical syntax.

The SSVD does not include our entire collection, nor a comprehensive list of all of the documentation available for our studies. Included in the variable descriptions.

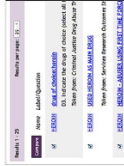
college major

GO

- Browse by Study
- Browse by Series

## Comparing Variables

Clicking 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 label.

Proportional election will return only variables starting with all proportional elections.

Proportional election Obama will return only variables starting with the year 2008 and including Obama.

Use quotes to search for specific phrases.

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

11% satisfaction

# ICPSR: Compare Variables

## Compare Variables

COMPONENT	CHARACTER	BACK TO SUMMARY TABLE
MAJOR1	MAJOR2	OC12
General Social Surveys, 1972-2014 Excluded: 1974	General Social Surveys, 2012-Merged Data, Excluded: 2007-2011; Low Marginal Content File Data File	Longitudinal Study of American Youth, Excluded: 2007-2011; Low Marginal Content File Data File
COLLEGE MAJOR 1	COLLEGE MAJOR 2	COLLEGE MAJOR
<p><b>Question</b></p> <p>What was your major or field of study (RESPONDENT'S COLLEGE DEGREE) DEGREE? IF THE RESPONDENT RECEIVED MORE THAN ONE DEGREE, ASK ABOUT THE HIGHEST DEGREE ABOUT THE HIGHEST DEGREE OBTAINED UP TO TWO MENTIONS.</p>	<p><b>Question</b></p> <p>What was your major or field of study (RESPONDENT'S COLLEGE DEGREE) DEGREE? IF THE RESPONDENT RECEIVED MORE THAN ONE DEGREE, ASK ABOUT THE HIGHEST DEGREE ABOUT THE HIGHEST DEGREE OBTAINED UP TO TWO MENTIONS.</p>	<p><b>Question</b></p> <p>COLLEGE MAJOR</p>
<p><b>Responses</b></p> <p>VALUE LABEL 0 14P</p>	<p><b>Responses</b></p> <p>VALUE LABEL 0 14P</p>	<p><b>Responses</b></p> <p>VALUE LABEL .00 Not a student</p>
<p><b>Relationships</b></p> <p>0 14P</p>	<p><b>Relationships</b></p> <p>0 14P</p>	<p><b>Relationships</b></p> <p>0 14P</p>

Psych M144/Stats M154 Fall 2017

19 / 30

# ICPSR: Study Page and Data

**ICPSR 38423**

High School Longitudinal Study  
2009-2013 (United States) (ICPSR 38423)

**Funded by:**  
Private funding by state.

**Principal Investigator:**  
James Heckman

**Principal Investigator's Department:**  
Department of Economics, University of California, Berkeley

**Abstract:**  
This study is a part of the High School Longitudinal Study (HSLS), a national longitudinal study of high school students in the United States. The study follows a nationally representative sample of high school students from the 2009-2010 school year through the 2012-2013 school year. The study includes data on students' academic achievement, social skills, and other factors. The study is designed to provide information on the factors that influence students' educational and career outcomes.

**Access Notes:**  
ICPSR 38423 is available for use by registered users of the ICPSR database. For more information, please contact the ICPSR staff.

**Characteristics:**  
ICPSR 38423 is a longitudinal study of high school students in the United States. The study includes data on students' academic achievement, social skills, and other factors. The study is designed to provide information on the factors that influence students' educational and career outcomes.

**ICPSR Study Page**

**Study Description**

**Creation:**

Psych M144/Stats M154 Fall 2017

20 / 30

# ICPSR: Reading About Data

FIN 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 bibliography of data-related literature. The *Bibliography of Data-Related Literature* is a searchable bibliography of data-related literature. It contains over 70,000 citations of published and unpublished works resulting from data-related research held in the ICPSR archive. Developed with support from the National Science Foundation (ISE-9377984), the *Bibliography* represents over 50 years of data-related research.

## Data-Related Publications

The ICPSR *Bibliography of Data-Related Literature* is a continuously updated bibliography of data-related literature. The *Bibliography of Data-Related Literature* is a searchable bibliography of data-related literature. It contains over 70,000 citations of published and unpublished works resulting from data-related research held in the ICPSR archive. Developed with support from the National Science Foundation (ISE-9377984), the *Bibliography* represents over 50 years of data-related research.


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

GO

- View all citations
- Browse by author
- Browse by journal
- Browse by study

## Search Tips

Click on the links in this...



Psych M144/Stats M154 Fall 2017

21 / 30

# ICPSR: Publications linked to Data

FIN DATA SEARCH COMPARE VARIABLES DATA-RELATED PUBLICATIONS RESOURCES FOR STUDENTS HELP

## Filter Results

Pub. Year \*

Author \*

## Search Results

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

Basics tab selected

optimize precision mental health

REFINED SEARCH

Your query returned 1 citations.

Sort by: Relevance

Export	Clear All	Pub. Year	Citation
<input type="checkbox"/>	<input type="checkbox"/>	1992	Romm, Robert, Scherer, Michael F., Bergeman, C.S., Pedersen, M.L., et al. <i>Personality and Individual Differences: A Meta-Analysis of the Relationship Between Intelligence and Personality</i> . <i>Personality and Individual Differences</i> , 13, (6), 927-930. Full Text Options: DOI # WorldCat # Google Scholar # Export Options: RIS EndNote XML

Results 1 - 1

Psych M144/Stats M154 Fall 2017

22 / 30

## 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 Literatures:

Bloomin, Robert, Scheider, Michael E., Bergeman, C.S., Pedersen, N.L., Nesselroade, J.B., McClearn, G.E., Aug 1992. **Opt pessimism and mental health: A twin/adoption analysis.** *Personality and Individual Differences* 13, (6): 921-930.  
doi:10.1016/0191-8869(92)90099-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

Psych M144/Stats M154 Fall 2017

23 / 30

## Data Sources

- ICPSR: <http://www.icpsr.umich.edu/icpsrweb/ICPSR/>
- UCLA Library Data Archive Catalog: [https://dataverse.harvard.edu/datasetverse/ssda\\_ucla](https://dataverse.harvard.edu/datasetverse/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](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/>

Psych M144/Stats M154 Fall 2017

24 / 30

## 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 with Data Management Plans

- Data Archive
- ICPSR

## Questions?



Psych M144/Stats M154 Fall 2017

29 / 30

## 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

Psych M144/Stats M154 Fall 2017

30 / 30