

Survey Search

...

CS 30122 Project

<https://github.com/bethbailey/survey-search>

Group members

Bethany Bailey

Leoson Hoay

Ruixue Li

Group Name

Pawsitive (from previous project idea “Catalog”)

Project description/goals

Build an interactive social science survey data repository where students and researchers can search surveys by keyword and topic on multiple levels (e.g. search for entire surveys and individual questions).

Include the ability for users to upload their own survey data to the repository to share with other users. User-uploaded surveys will be automatically processed to enable keyword and topic search.

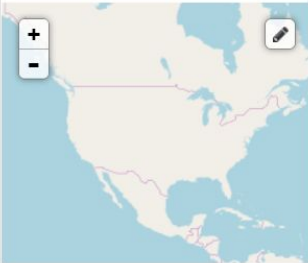
Order by:
Popular

Datasets ordered by Popular

Filter by location

Clear

Enter location...



Map tiles & Data by [OpenStreetMap](#), under [CC BY SA](#).

231,277 datasets found

USCIS My Case Status 1179 recent views

Department of Homeland Security – USCIS provide a way for the public who applied for U.S. citizenship to check the status of their application online. To view the status of a case, the application...

HTML

ZIP Code Data 999 recent views

Department of the Treasury – This study provides detailed tabulations of individual income tax return data at the state and ZIP code level.

HTML

Federal

Federal

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

General Social Survey Data Explorer

[About](#) [Contact](#) [Register](#) [Sign In](#)

GSS Data Explorer

[Features](#) [Media Room](#) [Help and Resources](#) [Search GSS Variables](#) [Key Trends](#)

MY GSS > Search Data

3 VARIABLES

Search Data

Q Keyword

All

Enter a Keyword

Filter by Years: 1972 - 2016

or Select Specific Years

Filter by Module / Subject:

Choose Modules / Subjects

+

SEARCH

5867 Results Matching Criteria

Collapsed View

Expanded View

+ ALL	Variable Name	Description	Years Available
✓	year	Gss year for this respondent	1972 - ----- 2016
✓	id	Respondent id number	1972 - ----- 2016
+	wrkstat	Labor force status	1972 - ----- 2016

Data

Source

- Social science surveys such as the General Social Survey
 - Potential sources:
 - General Social Survey, British Social Attitudes Survey, European Social Survey, etc.

Collection method

- Download datasets from official websites and/or scrape survey websites
- Allow users to potentially upload their datasets

Expected finishing date for data collection

- 6th week

What needs to be done and task distribution

1. Gather data from various sources (Everyone)
2. Clean data (Bethany)
3. Build web interface (Ruixue and Leoson)
4. Topic modeling (Ruixue and Bethany)
5. Search algorithm (Bethany and Leoson)
6. End-User Feedback

Potential Issues

- Issues with data collection: many large scale surveys are not open to the public
 - Potential Solutions
 - Use many smaller surveys/ allow user input
 - Talk to faculty and survey organizations and ask for permission
 - Scrape survey data from web
 - Expand to include other types of datasets
- Difficulties with multiple data formats for user input
 - Potential solution
 - Create clear guidelines for the format of the data input
 - Automatically check for compatibility

Questions/Comments?