1. Indoor Scene Recognition

http://web.mit.edu/torralba/www/indoor.html

The database contains 67 Indoor categories, and a total of 15620 images. The number of images varies across categories, but there are at least 100 images per category. All images are in jpg format. The images provided here are for research purposes only.



2. MovieLens

https://grouplens.org/datasets/movielens/

This dataset (ml-20m) describes 5-star rating and free-text tagging activity from MovieLens, a movie recommendation service. It contains 20000263 ratings and 465564 tag applications across 27278 movies. These data were created by 138493 users between January 09, 1995 and March 31, 2015. This dataset was generated on October 17, 2016.

This dataset provides users rating, movie's tag selected by users, movie's own tag, movie's title and type.

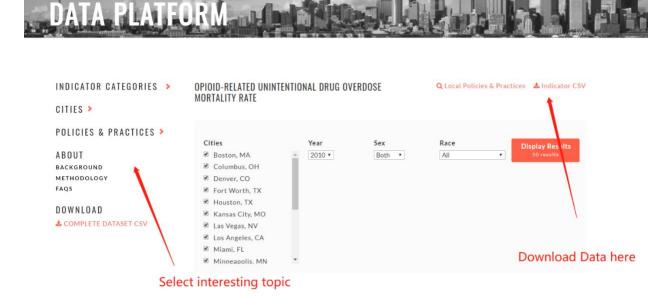
Furthermore, this dataset provides IMDB ID for each movie. By usage the IMDB ID with python library IMDbPY https://imdbpy.github.io/ You can grab every movie's detail information from IMDB, like Company, actor, director, IMDB rating and so on.

3. Big Cities Health Inventory Data

https://bchi.bigcitieshealth.org/indicators/1892/searches/34448

The Health Inventory Data Platform is an open data platform that allows users to access and analyze health data from 26 cities, for 34 health indicators, and across six demographic indicators.

Include many useful health data of different country in different year. You can combine different type of data to find out the potential relationship behind this data. For example: combine Adult smoking and Teen smoking data with Lung cancer data.



For some code appear in the data set, you can refer to this page: https://www.bigcitieshealth.org/methodology

For example, in cancer data C50 means female breast cancer and C33-C34 means lung cancer.

4. American Housing Survey (Metropolitan)

https://www.census.gov/programs-surveys/ahs/data.2017.html

// Census.gov > Our Surveys & Programs > American Housing Survey (AHS) > Data

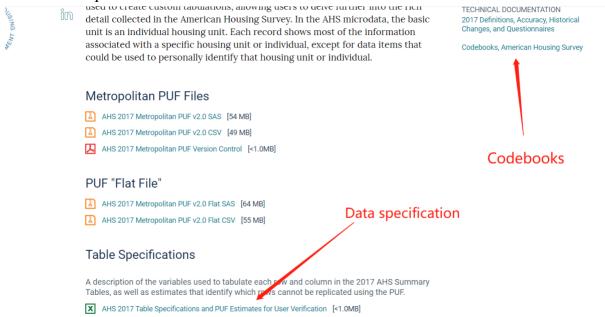
Microdata are files containing individual responses to survey questions. They are used to create custom tabulations, allowing users to delve further into the rich detail collected in the American Housing Survey. In the AHS microdata, the basic unit is an individual housing unit.

Each record shows most of the information associated with a specific housing unit or individual, except for data items that could be used to personally identify that housing unit or individual.

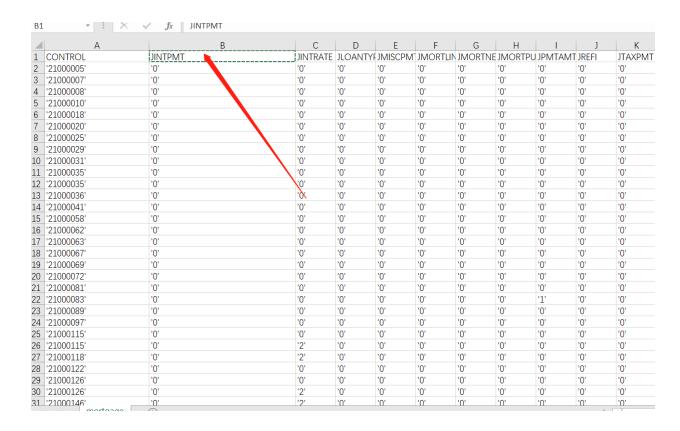


To find out what's the meaning of the data, you can refer to the codebooks: https://www.census.gov/data-tools/demo/codebook/ahs/ahsdict.html

And Table Specifications:



For example, in mortgage.csv, you may wonder what's the meaning of "JINTPMT":



You can either check the description in the Table specification file:

TABLE ROW/BOXHEAD SPECIFICATIONS:							Cells in C	Green are tabl
Housing Unit Chatracteristics	<u>C-00</u>	household	X				either in the	e column to t
General Housing Data	<u>C-01</u>	household	X	X	X	X		
Rooms, Size, and Amenities	C-02	household	X	X	X	X		
Heating, Air Conditioning, and Appliances	C-03	household	X	X	X	X		
Plumbing, Water, and Sewage Disposal	C-04	household		X	X	X		
Housing Problems	C-05	household		X	X	X		
Housing Migration—Previous Unit	C-06	household		X	X	X		
Housing and Neighborhood Search and Satisfaction	C-07	household		X	X	X		
Household Demographics	C-08A	household		X	X	X		
Disabilities	C-08B	household		X	X	X		
Income Characteristics	C-09	household		X	X	X		
Housing Costs	C-10	household		X	X	X		
Value, Purchase Price, and Source of Down Payment	C-13	household			X			
Mortgage Characteristics		household &						
	C-14A	mortgage	ck this		Х			
Additional Mortgage Characteristics	C-14B	mortgage			X			
Home Improvement Characteristics	<u>C-15</u>	household			X			
Home Improvement Costs	C-16	projects			X			
Rent Subsidies and Rental Property Management	<u>C-17</u>	household				X		
Neighborhoods	S-03/C-18	household		X	X	X		
Emergency and Disaster Preparedness	S-06	household		X	X	X		
Delinquent Payments and Notices	S-08	household		X	X	X		
Evictions	S-12	household		X	X	X		
Commuting to Work by Mode of Transportation	S-13A	household		Х	X	X		
Annual Commuting Costs by Type of Commuter	S-13B	household		X	X	X		
C=Core								
S=Special Topic/Supplemental								

Or search "INTPMT" (without prefix J) in Codebook:

