FINDING QUANTITATIVE DATASETS

A dataset is a collection of data. Quantitative datasets contain data that can be quantified, or measured. Usually a quantitative dataset refers to the contents of a data table, in which each column represents a variable, and each row an item in the dataset.

Tips for getting started:

☐ **Define your goals.** Who or what are you trying to study? What time period or range are you interested in? What geographic region or location do you need?

☐ **Consider who might collect the data you need.** For example, government agencies, trade groups, commercial institutions, academic researchers.

Places to search for data:

○ See if the libguide for your discipline has listed any data sources. You can get to the guides from the Library homepage or at: guides.boisestate.edu.

○ Try a data repository. You can find a variety of repositories listed by subject matter at re3data.org.

○ Use Google Dataset Search - datasetsearch.research.google.com.

DATA VERSUS STATISTICS

In regular conversation, the words "data" and "statistics" are often used interchangably. In research, a distinction is made between the two. Data are the raw material from which statistics are produced. Or, in reverse, statistics summarize data. - Michigan State University Libraries

Source: Statistical Abstracts of the U.S.

Table 1007. Subsidized Apartments Completed by Type of Subsidy: 2015 to 2019					
Completions and type of subsidy	2015	2016	2017	2018	2019
Total apartments completed............	310,000	310,100	346,800	335,600	342,900
Subsidized apartments completed..............	29,160	20,160	35,200	31,910	36,840
Percent subsidized..............	9.4	6.5	10.1	9.5	10.7
Subsidized by type: [1]					
Section 8............	10,680	6,010	12,080	16,570	10,090
Housing for Elderly Direct Loan Program...........	1,163	1,298	(S)	1,790	1,536
Low Income Housing Tax Credit (LHITC)............	19,190	13,310	21,630	15,020	24,930
Federal tax exempt multifamily bond financing.....	(S)	1,003	2,223	(S)	2,870
Other subsidized programs...............	6,881	6,442	10,080	9,871	10,500

Statistics answer the question "how much?" or "how many?" They often come in the form of a statistical table, such as in the example on the left.

Source: MILES Dataset 25

Data are collected and analyzed in order to investigate a phenomenon. Raw data is often (but not always) in the form of a dataset that can be analyzed using software such as SPSS. To the right is an example of a data table.

	A	B	C	D	E	F	G	H
1	ID	ID_#	AVG_TEMP	AVG_WIND	CORT_CONA	CORT_CONB	TEST_CONA	TEST_CONB
2	1	O-O-1	4.5	4.47				
3	2	O-O-2	6.8	3.73	0.057	0.297	63.335	95.272
4	3	O-O-3	6.8	3.73	0.023	0.082	109.817	148.287
5	4	O-O-4	6.8	3.73	0.104	0.291	72.215	103.839
6	5	O-O-5	3.2	3.17				
7	6	O-O-6	3.2	3.17				
8	7	O-O-7	4.6	1.94	0.106	0.077	189.997	139.721
9	8	O-O-8	4.95	1.97	0.22	0.243	154.131	159.975
10	9	O-O-9	4.82	2.37	0.337	0.045	106.321	80.884
11	10	O-O-10	4.82	2.37	0.495	0.146	69.338	29.736

For more information and to learn about how the Library's Research Data Management Team can help, visit: https://www.boisestate.edu/library/research-support/finding-data