# 2. WORKING WITH THE CENSUS DATA API

The U.S. Census Bureau has produced an API User Guide and organized a Webinar to help developers and researchers access and use the Census Data API to request data from Census Bureau data sets.<sup>18</sup> Key information from these resources is summarized below.

## **API Key**

Any user may query small quantities of data with minimal restrictions (up to 50 variables in a single query, and up to 500 queries per IP address per day). However, more than 500 queries per IP address per day requires that you register for an API key.

To request an API key:

• Go to <www.census.gov/developers/>.

- Click on the Request a KEY box on the left side of the page.
- Fill out the form: <https://api.census.gov/data /key\_signup.html>.
- You will receive an email with your key code and activation instructions in the message.

Once you have an API key, you can extract information from Census Bureau data sets using a variety of tools including JSON, R, Python, or even by typing a query string into the URL of a Web browser.<sup>19</sup>

### Components of an API Query

Each API call, or query, requires a series of components to function properly. Figure 2.1 provides an example of an API call and its components.

<sup>19</sup> Users may not be able to view the results from API queries in all Web browsers, but Firefox and Chrome provide this functionality.



<sup>&</sup>lt;sup>18</sup> U.S. Census Bureau, Developers, Census Data API User Guide, <www.census.gov/data/developers/guidance/api-user-guide.html>; American Community Survey (ACS), Using the Census API with the American Community Survey Webinar, <www.census.gov/data /academy/webinars/2019/api-acs.html>.

With the API, you access only the variables and geographic areas that you need. In the query above:

- The "Census Data API" specifies the API that is being used to access the data.
- The "Dataset" specifies the data source is the 2016 ACS 1-Year Detailed Tables.
- The "Get Function" (get=) specifies the variable(s) you are requesting the API to give you.
- The "Variable List" includes the variable(s) you are requesting. You can include up to 50 variables in a single API query (separated by commas). In this data set, the variable called NAME provides the name of the geographic area(s) that you are using to limit your search.
- The "Predicate" clause specifies how variables should be filtered or limited (for example, for certain geographic areas).
- "Geography" specifies the geographic area(s) of interest.

### Understanding Variable Names

Each variable in a data set has a name, which may have meaning on its own (for example, TRACT for census tract, or NAME for geographic area name) or may be an alpha-numeric identifier. In the American Community Survey (ACS), many of the variable names are alpha-numeric, such as B01003\_001E, which is the Total Population. The "Variable List" column on the Census API Datasets Web page provides links to all the variables in each data set.<sup>20</sup>

The first letter in an ACS variable's name indicates the table type. For example, a "B" at the beginning of a name indicates that data are from a "base" table and "C" is for a collapsed table. The collapsed tables cover the same topics as the base tables, but with fewer details.<sup>21</sup>

The next five digits in an ACS variable name refer to the rest of the summary table identifier (ID). The first two digits are a subject identifier and the next three are a sequential number. Tables beginning with "01," for example, are for age and sex, "08" tables are commuting (journey to work) and place of work.<sup>22</sup> The next three digits reflect the table number within a subject.

Some variables end in "A" through "I," which tells you that the corresponding ACS table provides characteristics that are repeated for different race and Hispanic origin groups. For example, table numbers ending in a "C" are for American Indian and Alaska Native Alone populations. Table numbers with an "H" suffix are for non-Hispanic White populations. For example, Table B01001H is Sex by Age (White Alone, Not Hispanic or Latino).

Other tables end in "PR," which tells you that the data came from the Puerto Rico Community Survey instead of the ACS. These Puerto Rico-specific tables exist because the wording of the Puerto Rico Community Survey questionnaire for some subjects differs slightly from the ACS questionnaire.

The six-character table ID is followed by an underscore and three more digits. Those three digits refer to the line number within a table. For example, "001" may refer to the total, "002" may refer to males, and so on.

Finally, the last character in an ACS variable is an alphabetical suffix (E, M, PE, or PM).

- "E" refers to a numeric representation of the ACS estimate.
- "M" refers to a numeric representation of the margin of error.
- "PE" refers to an estimate representing a percent of the total.
- "PM" refers to the margin of error for a percentage.

In some data sets, users may also see variables ending with "EA," "MA," "PEA," "PMA," or "SS." These suffixes are special annotations used to communicate information about estimates, margins of error, or statistical significance. For example, "SS" refers to "Statistical Significance" and is only included in the Comparison Profile tables. When extracting data for ACS estimates or margins of error, it is important to also extract the data for any special annotations.

 $<sup>^{\</sup>rm 20}$  U.S. Census Bureau, Census API: Datasets in /data and its descendants, <a href="https://api.census.gov/data.html">https://api.census.gov/data.html</a>.

<sup>&</sup>lt;sup>21</sup> Detailed information about ACS table IDs is available on the Census Bureau's Table IDs Explained Web page at <www.census.gov /programs-surveys/acs/guidance/which-data-tool/table-ids-explained .html>.

<sup>&</sup>lt;sup>22</sup> Data Profiles, Narrative Profiles, Comparison Profiles, and Selected Population Profiles cover multiple topics, so they do not have any characters to indicate a subject.

Suppose you needed an estimate of the male population aged 5 to 9. Those data are located in Table B01001: Sex by Age. Data for males aged 5 to 9 appear within that table on line 4. Finally, estimates are designated by an "E." Thus, the variable string to include in your API query would be B01001\_004E (see Figure 2.2). For more information about variable formats, types, and annotations, see the Census Bureau's Notes on ACS 5-Year Data.<sup>23</sup>

<sup>23</sup> U.S. Census Bureau, Developers, Notes on ACS 5-Year Data, <www.census.gov/data/developers/data-sets/acs-5year/data-notes .html>.



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/>.

## Filtering Geography

In an API query, you can use a "predicate" to filter your ACS results by geography:

- The "for" predicate (&for) restricts the variables by geography at various levels.
- The "in" predicate (&in) restricts the geographic scope and can be used in combination with a "+" sign to further specify a geographic area of interest.

An asterisk can be included as a wildcard to search for all the values of a geographic area or a string variable; however, you cannot build a predicate with wildcards for numeric variables. Table 2.1 provides several examples of predicates of geography you can use when building your queries.

Every query must include a geography. The Census Data API supports Federal Information Processing Series (FIPS) codes and Geographic Names Informational System (GNIS) codes. You may look up codes for certain geographic areas on the Census Bureau's Geography Web page.<sup>24</sup> You can also easily find specific codes by using the wildcard with a geographic level of interest in the API.

<sup>&</sup>lt;sup>24</sup> U.S. Census Bureau, Geography, American National Standards Institute (ANSI) Codes, <www.census.gov/library/reference/code-lists /ansi.html>.

Table 2.1. Examples of Predicates for Geography					
Predicate	Action				
&for=state:*	Retrieves the result for all states.				
&for=state:01	Restricts the result to include only Alabama (state code = 01).				
&for=county:*∈=state:01	Restricts the result to include all counties in Alabama.				
&for=county:001∈=state:01	Restricts the result to include only Autauga County (County: 001), Alabama.				
&for=county (or part):*∈=state:01+place:62328	Restricts the result to include all counties (or portions of coun- ties) within Prattville city (Place: 62328), Alabama.				
&for=county (or part):073∈=state:01+place:07000	Restricts the result to include the portion of Jefferson County (County:073), Alabama that is within Birmingham city (Place:07000).				
Source: U.S. Consus Rureau					

For example, to retrieve data from Table B01001\_001E (total population) for all states, use the following the string:

<https://api.census.gov/data/2016/acs/acs1?get=NAME ,B01001\_001E&for=state:\*>

The results of this query are displayed in Figure 2.3.

Figure 2.3. Results of Query for Total Population by State: 2016						
https://api.census.gov/data/2016 × +						
← → C ☆ https://api.census.gov/data/2016/acs/acs1?get=NAME,B01001_001E&for=state:*						
<pre>[["NAME","B01001_001E","state"], ["Alabama","4863300","01"], ["Alaska","741894","02"], ["Arkansas","2988248","05"], ["California","39250017","06"], ["Colorado","5540545","08"], ["Connecticut","3576452","09"], ["Delaware","952065","10"], ["Delaware","952065","10"], ["Florida","20612439","12"], ["Florida","20612439","12"], ["Georgia","10310371","13"], ["Hawaii","1428557","15"], ["Idaho","1683140","16"], ["Illinois","12801539","17"], ["Indiana","6633053","18"], ["Iowa","3134693","19"], ["Kansas","2907289","20"], ["Kentucky","4436974","21"],</pre>						
Source: U.S. Census Bureau, <https: 2016="" acs="" acs1?get="NAME,B01001_001E&amp;for=state:*" api.census.gov="" data="">.</https:>						

The results shown in Figure 2.3 are in JavaScript Object Notation (JSON) format. JSON, like XML, is a simple format for exchanging data between platforms using human-readable language. In order to return results in a concise manner, the Census Bureau uses a nonstandard, streamlined version of JSON:

- Data are represented in a two-dimensional array.
- Square brackets [ ] hold arrays.

- Values are separated by commas (,).
- The first line of data contains the variable names.
- Each subsequent line of data is a record for a given geography.

Data users familiar with JSON can convert results into a standard JSON format using the following code snippet:

```
```js
let standardJSON =
    <fetch stats async with promise>
    .then(data => {
        let labels = data[0].map(datum => datum.toUpperCase());
        let rows = data.slice(1);
        let objArray = rows.map(row => { return Object.assign(
            {},
            ...labels.map((key, idx) => ({ [key]: row[idx] }))
        );
        });
```

To query 2016 total population only for Alabama (01), California (06), and Rhode Island (44), we could use the following string:

In this example, the predicate "for" is used to limit the search to Alabama (state code = 01), California (state code = 06), and Rhode Island (state code = 44). (See Figure 2.4).

<https://api.census.gov/data/2016/acs/acs1?get=NAME ,B01001\_001E&for=state:01,06,44>.

Figure 2.4. Results of Query for Total Population of Selected States: 2016						
https://ap	pi.census.gov/da ×					
$\textbf{\leftarrow} \ \Rightarrow \ \textbf{G}$	Secure   https://api.census.gov/data/2016/acs/acs1?get=NAME,B01001_001E&for=state:01,06,44					
[["NAME","B0 ["Alabama"," ["California ["Rhode Isla	1001_001E","state"], '4863300","01"], *","39250017","06"], and","1056426","44"]]					
Source: U.S. Cer	isus Bureau, <https: 2016="" acs="" acs1?get="NAME,B01001_001E&amp;for=state:01,06,44" api.census.gov="" data="">.</https:>					

To query 2016 total population for counties included in the 2016 ACS 1-year data set for California (06), we could use the following string:

<https://api.census.gov/data/2016/acs/acs1?get=NAME ,B01001\_001E&for=county&in=state:06>.

The predicate "for" limits the search to counties, and the "in" term limits the search to California (&in=state:06). (See Figure 2.5).

Note that certain predicates will result in an error ("unknown/unsupported geography hierarchy") if the specified geographic area is not available in the data set. For example, estimates for small geographic areas with fewer than 20,000 people are only available through the ACS 5-year data.

Figure 2.5. Results of Query for Total Population of Counties in California: 2016					
https://api.census.gov/da ×					
← → C Secure   https://api.census.gov/data/2016/acs/acs1?get=NAME,B01001_001E&for=county∈=state:06					
<pre>[["NAME","B01001_001E","state","county"], ["Alameda County, California","1647704","06","001"], ["Butte County, California","226864","06","007"], ["Contra Costa County, California","1135127","06","013"], ["El Dorado County, California","135625","06","017"], ["Fresno County, California","136646","06","023"], ["Humboldt County, California","1368483","06","025"], ["Kern County, California","180883","06","025"], ["Kern County, California","180883","06","029"], ["Kings County, California","149785","06","031"], ["Lake County, California","164116","06","033"], ["Los Angeles County, California","154697","06","039"], ["Madera County, California","154697","06","041"], ["Marin County, California","87628","06","045"], ["Mendocino County, California","435232","06","047"], ["Monterey County, California","435232","06","053"],</pre>					
Source: U.S. Census Bureau, <https: 2016="" acs="" acs1?get="NAME,B01001_001E&amp;for=county&amp;in=state:06" api.census.gov="" data="">.</https:>					

### Groups

You can include up to 50 variables in a single API query, but it is possible to query more than 50 variables with a single call using the "groups" function. For example, to query all elements from Table B01001 (Sex by Age, for the United States) from the 2016 ACS 1-year data, you would use the following string:

<https://api.census.gov/data/2016/acs/acs1?get =group(B01001)&for=us:\*>. That query returns all elements of the Sex by Age Table, with variable names in the first line and data in subsequent lines (see Figure 2.6).

For more information about groups, visit the Census Bureau's Web page on Groups Functionality.<sup>25</sup>

<sup>25</sup> U.S. Census Bureau, Developers, Groups Functionality-Pull more than 50 variables with 1 call, <www.census.gov/data/developers /updates/groups-functionality.html>.

Source: U.S. Census Bureau, <https://api.census.gov/data/2016/acs/acs1?get=group(B01001)&for=us:\*>.

# Sample Query: Hmong Population by State

Suppose you needed information on the Hmong population living in each state. How would you use the Census Data API to get that information? There are many ACS data sets available through the Census Data API. For this example, we focus on the ACS 1-Year Detailed Tables because we need statelevel data for the most recent year.

Figure 2.7 provides an example of an API call and its components that could be used for this query.



Here are steps you can use to assemble the components of this query:

- 1. Start your query with the Census Data API Web address: "https://api.census.gov/data."
- Add the data set year (2016) and name acronym for the 1-Year Detailed Tables, and follow this base URL with a question mark: "https://api.census.gov/data/2016/acs/acs1?."
- 3. Add variables starting with a get clause, "get=": "https://api.census.gov/data/2016/acs/acs1?get=."
- Add the variable NAME, which provides the name of the geographic area(s) that you will use to limit your search: "https://api.census.gov/data/2016/acs /acs1?get=NAME."
- Find your variable of interest (B02015\_009E) by searching for "Hmong" among the list of variables for the 2016 ACS 1-Year Detailed Tables.<sup>26</sup>
- 6. Use a comma to separate the NAME and B02015\_009E variables. Add an additional

variable, B02015\_009M, to get the margin of error: "https://api.census.gov/data/2016/acs/acs1?get =NAME,B02015\_009E,B02015\_009M."

- Add geography using a predicate clause starting with an ampersand (&) to separate it from your "get" clause and then a "for=" to identify geographic areas of interest: "https://api.census.gov/data/2016/acs/acs1?get =NAME,B02016\_009E,B02015\_009M&for=."
- Identify the geographic area(s) that you need (state) by reviewing the list of geographies available for the 2016 ACS 1-Year Detailed Tables.<sup>27</sup>
- Because you need data for all the states, add an asterisk (after a colon) to get all state values: "https://api.census.gov/data/2016/acs/acs1?get =NAME,B02015\_009E,B02015\_009M&for=state:\*."

If you are using an API key, insert "&key=" followed by your key code at the end of your query.

<sup>&</sup>lt;sup>26</sup> U.S. Census Bureau, Census Data API: Variables in <https://api.census.gov/data/2016/acs/acs1/variables.html>.

<sup>&</sup>lt;sup>27</sup> U.S. Census Bureau, <https://api.census.gov/data/2016/acs/acs1/geography.html>.

The results of the query are shown in Figure 2.8. In this figure, some of the data cells for Alaska and Delaware include "null" values. This means that there are no data available for a particular query.

If your query returns an error message with no data, check your spelling, capitalization, and spacing. Correct any errors and run the query again.



Source: U.S. Census Bureau,<https://api.census.gov/data/2016/acs/acs1?get=NAME,B02015\_009E,B02015\_009M&for=state:\*>.

# *Getting Data for Detailed Population Groups*

Suppose you also wanted to get information on the percentage of Hmong Americans with a bachelor's degree or higher. You could use the Selected Population Profile (SPP) data set to find this information for the nation and for states with relatively large populations of Hmong Americans.

The first step is to find a code that we can use to identify the Hmong subpopulation in the SPP data set. Here is a sample query:

<https://api.census.gov/data/2016/acs/acs1/spp ?get=S0201\_0006E,POPGROUP,POPGROUP\_ TTL,NAME&for=us:\*>. In this query:

- The "spp?" identifier is included to query results from the Selected Population Profile data set.
- S0201\_0006E is the variable for total population, POPGROUP is the code for the population group, and POPGROUP\_TTL is the name of the population group, as shown in the list of variables for the 2016 ACS 1-Year Selected Population Profile.<sup>28</sup>

By scrolling through the results of this query, we find that the POPGROUP code for the Hmong alone population is "020" (see Figure 2.9).

<sup>28</sup> U.S. Census Bureau, Census Data API: Variables in <a href="https://api.census.gov/data/2016/acs/acs1/spp/variables.html">https://api.census.gov/data/2016/acs/acs1/spp/variables.html</a>.

#### Figure 2.9. Results of API Query for Population Groups: 2016

```
["17556935","012","Asian alone (400-499)","United States","1"],
["3813407","013","Asian Indian alone (400-401)","United States","1"],
["175592","014","Bangladeshi alone (402)","United States","1"],
["259554","015","Cambodian alone (405-409)","United States","1"],
["4214856","016","Chinese alone (410-419)","United States","1"],
["4048973","017","Chinese (except Taiwanese) alone (410-411)","United States","1"],
["165883","018","Taiwanese alone (412-419)","United States","1"],
["165883","018","Taiwanese alone (420-421)","United States","1"],
["2811885","019","Filipino alone (420-421)","United States","1"],
["278871","020","Hmong alone (422","United States","1"],
["76516","021","Indonesian alone (423-429)","United States","1"],
["769830","022","Japanese alone (430-439)","United States","1"],
["1438915","023","Korean alone (440-441)","United States","1"],
["205131","024","Laotian alone (442)","United States","1"],
["460515","026","Pakistani alone (447-449)","United States","1"],
["202744","028","Thai alone (447-449)","United States","1"],
["1803575","029","Vietnamese alone (450-459)","United States","1"],
```

Note: Data are only displayed for selected rows.

Source: U.S. Census Bureau,<https://api.census.gov/data/2016/acs/acs1/spp?get=S0201\_0006E,POPGROUP,POPGROUP\_TTL,NAME&for=us:\*>.

Now that we know the code for the Hmong population, we can get data on the share of Hmong Americans with a bachelor's degree or higher nationwide:

<https://api.census.gov/data/2016/acs/acs1/spp?get =S0201\_0099E,POPGROUP,POPGROUP\_TTL,NAME&f or=us:\*&POPGROUP=020>.

Here is a similar query for states:

<https://api.census.gov/data/2016/acs/acs1/spp?get =S0201\_0099E,POPGROUP,POPGROUP\_TTL,NAME &for=state:\*&POPGROUP=020>. The results from the U.S. query show that 18 percent of Hmong Americans aged 25 and older had a bachelor's degree or higher in 2016. Because of the relatively small Hmong population in many states, state-level SPP data on the Hmong population are only available for California and Minnesota (see Figure 2.10).

You can find other examples of queries for this data set by clicking the link in the "Examples" column on the API "datasets" page, which will take you to the examples page here: <a href="https://api.census.gov/data/2016/acs/acs1/spp/examples.html">https://api.census.gov/data/2016/acs/acs1/spp/examples.html</a>>.

## Figure 2.10. Results of API Query for the Share of Hmong Americans With a Bachelor's Degree or Higher: 2016 [["S0201\_0090E","S0201\_0099E","POPGROUP","POPGROUP\_TTL","NAME","POPGROUP","state"], ["48883","17.1","020","Hmong alone (422)","California","020","06"], ["34573","19.4","020","Hmong alone (422)","Minnesota","020","27"]]

Source: U.S. Census Bureau, <https://api.census.gov/data/2016/acs/acs1/spp?get=S0201\_0090E,S0201 \_0099E,POPGROUP,POPGROUP\_TTL,NAME&for=state:\*&POPGROUP=020>.

## **Opening Query Results in a Spreadsheet**

From your browser, you can save query results as a .csv file to work with the file in other programs. Here are the steps to open your query results in a spreadsheet:

1. Right click on the page and select "Save As" (see Figure 2.11).

https://api.census.gov/data/2016 × +		
← → C ☆ 🔒 https://api.census.gov/data/2016/ac	s/acs1?get=NAME,B02015_00	09E,B02015_009M&for=sta
["NAME","B02015_009E","B02015_009M","state"], "Alabama","0","195","01"], "Alaska",null,null,"02"],	Back	Alt+Left Arrow
"Arizona","201","201","04"], "Arkansas","1012","735","05"], "California","09720","0910","06"]	Reload	Ctrl+R
California, 985/2, 9219, 00 ], "Colorado", "5454", "2381", "08"], "Connecticut" "127" "160" "08"]	Save as	Ctrl+S
"Delaware", null, null, "10"],	Print	Ctrl+P
"District of Columbia","0","203","11"], "Elopida" "523" "562" "12"]	Cast	
"Georgia", "4619", "2119", "13"],	Translate to English	
"Hawaii","0","185","15"], "Idaho",null,null,"16"],	Adobe Acrobat	•
"Illinois", "1290", "1115", "17" ,	20	C+1.11

2. Save the page as a Comma-Separated Values (CSV) file by adding ".csv" at the end of the file name and selecting "All Files" as the file type (see Figure 2.12).

$\leftarrow$ $\rightarrow$ $\checkmark$ $\uparrow$ $\square$ $\Rightarrow$ This PC $\Rightarrow$	Documents > API examples	v ∿	Search API exam	nples 🔎
Organize 🔻 New folder				== • 🕐
	Name	Date modified	Туре	Size
File name: Hmong_Pop Save as type: All Files	_by_Stat <mark>e.csv</mark>			
∧ Hide Folders			Save	Cancel

3. Open the CSV file in a spreadsheet for further analysis of the query results (see Figure 2.13).

Figure 2.13. Opening the API Query Results								
H	•ూ లి - ∓						Hmong_Pop_by_	State - Excel
File	Home Insert Page La	yout Formulas I	Data Review View	v ACROI	BAT Q Tell	me what you wan	t to do	
A1 ▼ :: × ✓ & [["NAME"								
	A	В	С	D	Е	F	G	н
1	[["NAME"	B02015_009E	B02015_009M	state]				
2	["Alabama"	0	195	01]				
3	["Alaska"	null	null	02]				
4	["Arizona"	201	201	04]				
5	["Arkansas"	1012	735	05]				
6	["California"	98372	9219	06]				
7	["Colorado"	5454	2381	08]				
8	["Connecticut"	127	160	09]				
9	["Delaware"	null	null	10]				
10	["District of Columbia"	0	203	11]				
11	["Florida"	523	562	12]				
12	["Georgia"	4619	2119	13]				
Sour	rce: U.S. Census Bureau.							

The Census Bureau's Webinar on Using the Census API with the ACS provides more detail on how to open and manipulate your results in a spreadsheet.<sup>29</sup>

<sup>&</sup>lt;sup>29</sup> U.S. Census Bureau, American Community Survey (ACS), Using the Census API with the American Community Survey Webinar, <www.census.gov/programs-surveys/acs/guidance/training-presentations/acs-api.html>.