

Name _____ Period _____

Skill 14.1 Exercise 1

A researcher collects survey responses from 200 households, then uses a government report to compare the findings. Which is the primary data source, and which is the secondary data source?

Skill 14.2 Exercise 1

A temperature sensor records the following set of **raw readings**:

72°F, 69F, error, 21.1C, 70, seventy-one

Identify at least *three* reasons these data should be considered raw.

Convert the valid readings into a clean, consistent format in degrees Fahrenheit.

Skill 14.3 Exercise 1

Below is a list of different data items. For each one, identify **its data type** (e.g., Tabular, Non-Tabular, Image, Agnostic/Binary).

#	Data Item	Data Type
1	students.xlsx – contains student names, grades, and IDs	
2	report.txt – a text file of research notes	
3	diagram.png – a flowchart image of a process	
4	experiment.dat – a binary file storing sensor readings	
5	survey_results.csv – a table of survey responses	
6	photo.jpg – a landscape photograph	
7	metadata.xml – structured information about a dataset	
8	essay.rtf – a formatted written assignment	

Name _____ Period _____

Skill 14.4 Exercise 1

Open another browser tab, make an API call to request state-level data containing:

- The NAME variable,
- The total number of commuters, and
- The count for commuters who travel less than 5 minutes.

Find the appropriate variables to request in the [B08303 variable group documentation](#).

Use the example below as a guide for how to format this request,

`https://api.census.gov/data/2020/acs/acs5?get=NAME,B08303_001E&for=state:*`

Change this to
the correct value.

Copy and paste the API request along with the data for the first five states below.

API request:

Data for the first five states:

Make an API call to request state-level data containing:

- The NAME variable,
- The total commuters count, and
- The count for commuters who travel 90 or more minutes.

Copy and paste the API request along with the data for the first five states below.

API request:

Data for the first five states:

Name _____ Period _____

Skill 14.4 Exercise 2

Using the browser, make an API call to request state-level data containing the 'NAME' variable, the total commuters count, and the count for commuters who travel 90 or more minutes for only the state of New York.

Find the appropriate variables to request in the [B08303 variable group documentation](#).

Find the state code for New York here

https://en.wikipedia.org/wiki/Federal_Information_Processing_Standard_state_code#FIPS_state_codes

Copy and paste the API request along with the data.

API request:

Data:

Make an API call to request the same 3 variables for all counties within the state of New York.

Use the for parameter in conjunction with in to request specific geographic regions.

Copy and paste the API request along with the data for the first five counties.

API request:

Data for the first five counties:

Name _____

Period _____

Skill 14.6 Exercise 1

Indicate how you can import the requests module. Then, Make a GET request to the Census API to request county-level data containing,

- the NAME variable,
- the total commuters count, and
- the count for commuters who travel 90 or more minutes
- for all counties
- within the state of New York.

(see the previous exercise for this URL)

Store the response object in a variable called r.

Use the *text* attribute to access the returned data and store it in a variable called *r_text*. Then print the results.

Use the *json* method to access the returned data and store it in a variable called *r_json*. Then print the results

What is the advantage of using the *json* method?

Name _____ Period _____

Skill 14.7 Exercise 1

The get request below retrieves the total number of commuters for all states.

Do the following,

- Import the csv module
- Use the .json() method to access the decoded JSON data and store it in a variable called r_json.
- Write the JSON data into a CSV file called commute_data.csv

```
import requests

r =
requests.get('https://api.census.gov/data/2020/acs/acs5?get=NAME,B08303_001E&for=state:*)
```

Skill 14.8 Exercise 1

Below is a screenshot of a csv file with commuter data,

	NAME	B08303_001E	B08303_002E	state
1	Pennsylvania	5652158	179118	42
2	California	16710195	303289	06
3	West Virginia	697042	26827	54
4	Utah	1378826	51948	49

In the space below,

- Import the pandas library.
- Create a variable called commute_df. Set your variable to a pandas DataFrame from the commute_data.csv
- Rename the headings to more descriptive names. Below are descriptions for each code
 - B08303_001E = total commuters
 - B08303_002E = commuters with travel time under 5 minutes
- Preview the first few rows of commute_df using the *head()* method and print out the output