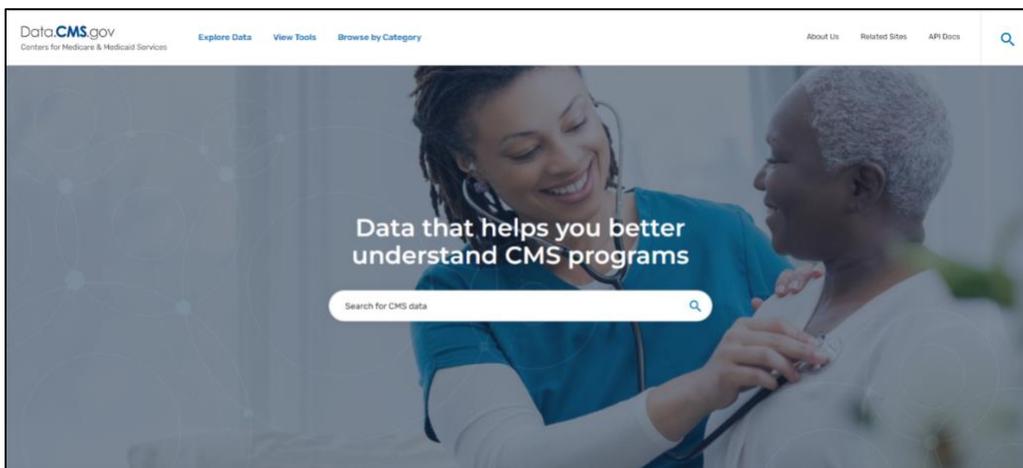# API FAQ for data.CMS.gov

MAY 2024 | VERSION 1.5

Prepared by:
The Centers for Medicare and Medicaid Services,
Office of Enterprise Data and Analytics

Data.CMS.gov
Centers for Medicare & Medicaid Services

Table of Contents

Data.**CMS**.gov
Centers for Medicare & Medicaid Services

Data.CMS.gov
Centers for Medicare & Medicaid Services

Revision History

Version	Date	Organization/Point of Contact	Description of Changes
1.0	2022-08-08	OEDA	Baseline Version
1.1	2022-12-06	OEDA	Adding Alt Text to Images
1.2	2022-12-16	OEDA	Functional FAQ Reference
1.3	2023-11-08	OEDA	Contact Email Update
1.4	2024-01-18	OEDA	API Download Documentation
1.5	2024-05-01	OEDA	Changed Example Dataset

Getting Started

Why doesn't my API call that worked before July 2021 work now?
Our platform completed a major update on July 21, 2021. API calls will need to be updated.

Do I need to register for an API key?
No, you do not currently need an API key or registration.

Do I need permission to use the data?
Some datasets may have license agreements restricting the use of the data but otherwise you may generally use the data on this site. Works of the U.S. Government are in the public domain and you don't need permission to reuse them, but an attribution to the agency as the source is appreciated. Your materials, however, shouldn't give the false impression that the government's endorsing your commercial products or services. See 42 U.S.C 1320b-10.

Finding Data

Is there a catalog of available data?
Yes, the data.json is an Open Data catalog containing the datasets available on our site. As new data is added, it is automatically updated to the data.json.

In the data.json, there is an array called **dataset**. You can search through this array using the title of your dataset, such as "Payroll Based Journal Daily Nurse Staffing." Inside of the dataset, there is an array called **distribution** which will contain all the versions of the data in all available formats for each version.

In this array there are a few different types of entries:

- The first entry (with a **description** of '**latest**') will provide a URL that will always point to the latest data, even as new versions are published.
- The remaining entries will provide references to the data at fixed points in time, listed in descending order by date. For example, https://data.cms.gov/data-api/v1/dataset/d10d792e-ea6e-4145-8512-34efbc1be04b/data will always point to the "Payroll Based Journal Daily Nurse Staffing" data from Q2 2021. Some datasets will have multiple historical versions available and some will only have the most recent data available.
    - The data with *mediaType* of "text/csv" is downloadable as a CSV file and data with *mediaType* of "application/zip" is downloadable as a ZIP file.
    - The data with a *format* of "API" will point to an API endpoint.

How can I get an earlier version of the data?
If your data has earlier versions available, there are two options to find the endpoint.

The first option is from the dataset page. There will be an "Access API" button at the top. Clicking on the button will launch a pop-up with a link to the API endpoint and a drop down with available time periods.

For example, on this page you can click the "Access API" button and select a year from the "Time Period" drop down.
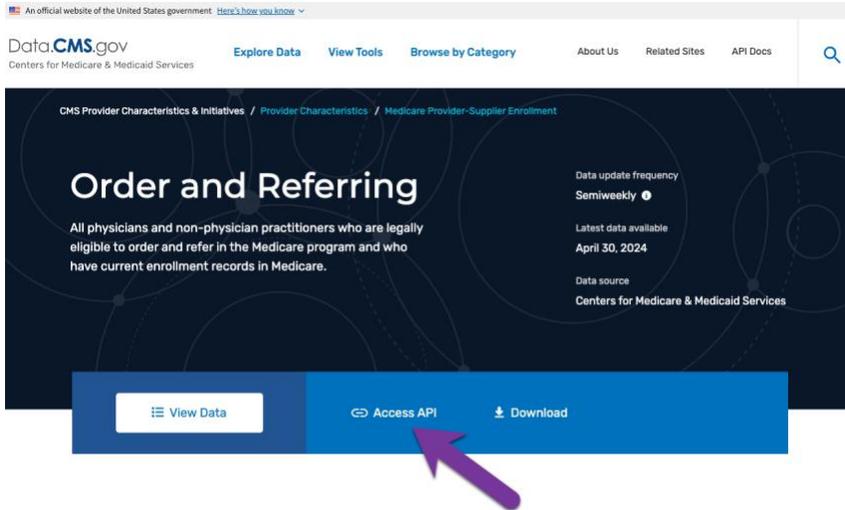
Data.**CMS**.gov
Centers for Medicare & Medicaid Services
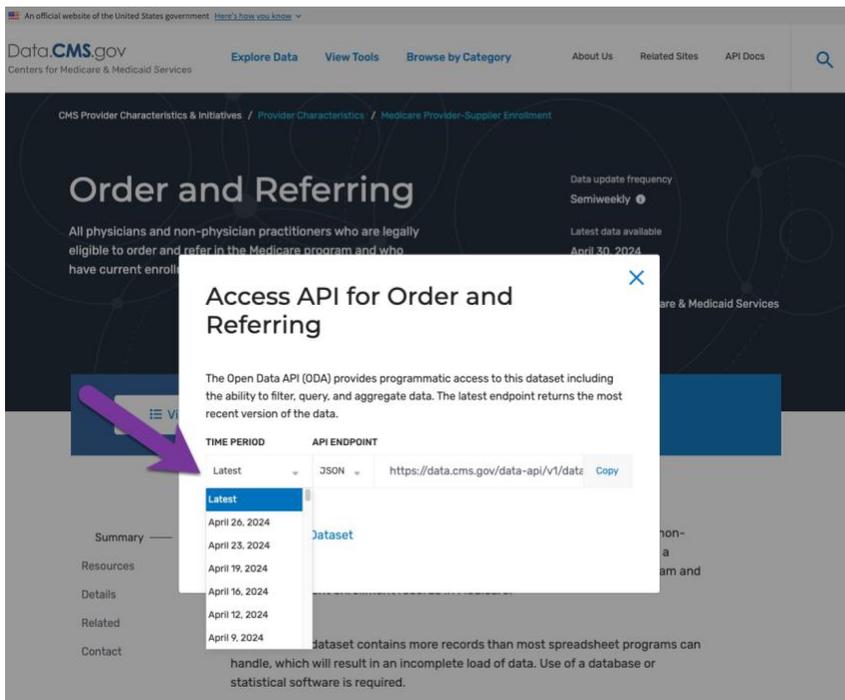
The second option is to use the data.json. There is a "temporal" field in the distribution section which will indicate the time period. An example is below.

Older Version Code Example

```
"""
Example in Python
```

Data.**CMS**.gov
Centers for Medicare & Medicaid Services

```
Finds the API link for the 2017 Medicare Inpatient Hospitals - by Provider
and Service Data
https://data.cms.gov/provider-summary-by-type-of-service/medicare-inpatient-
hospitals/medicare-inpatient-hospitals-by-provider-and-service
"""
import requests


url = "https://data.cms.gov/data.json"
title= "Medicare Inpatient Hospitals - by Provider and Service"
time_period = "2017-01-01/2017-12-31"

response = requests.request("GET", url)
if response.ok:
    response = response.json()
    dataset = response['dataset']
    for set in dataset:
        if title == set['title']:
            for distro in set['distribution']:
                if 'format' in distro.keys():
                    if distro['format'] == "API" and distro['temporal'] ==
time_period:
                        print(f"API link for {title} from {time_period} is
{distro['accessURL']}")
```

Is there a way to always get the latest data?

There are two ways to find the latest data.

Both methods will provide a URL that will point to the latest version of the data. The URL will not change as new versions are added. It is recommended to always start with the data.json rather than hardcoding any URL.

1. Use the distribution with the latest description:

   a. In the data.json the *dataset* elements will have a "*distribution*" array, in which the first element is marked with "*description*" of "*latest*". Find the "*accessURL*" element and it will have the URL to always points to the latest data.
   b. This URL returns data as a standard JSON.
   c. For example, this URL for Opt Out Affidavits is https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data

2. Use the URL in the identifier field:
   1. In the data.json the dataset will have a field called "*identifier"* and the value for this field is a URL that always points to the latest data.
   2. This URL returns data using the JSON:API specification, which has a different response structure that includes metadata about the dataset, otherwise it is identical to the URL in option 1.
   3. For example, this URL for Opt Out Affidavits https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data-viewer

Latest Data Code Example

```python
"""
Example in Python
Finds the API link for the latest Order and Referring Data
https://data.cms.gov/provider-characteristics/medicare-provider-supplier-
enrollment/order-and-referring
"""
import requests


url = "https://data.cms.gov/data.json"
title= "Order and Referring"

response = requests.request("GET", url)
if response.ok:
    response = response.json()
    dataset = response['dataset']
    for set in dataset:
        if title == set['title']:
            for distro in set['distribution']:
                if 'format' in distro.keys() and 'description' in
distro.keys():
                    if distro['format'] == "API" and distro['description'] ==
"latest":
                        latest_distro = distro['accessURL']
            print(f"The latest data for {title} can be found at
{latest_distro} or {set['identifier']}")
```
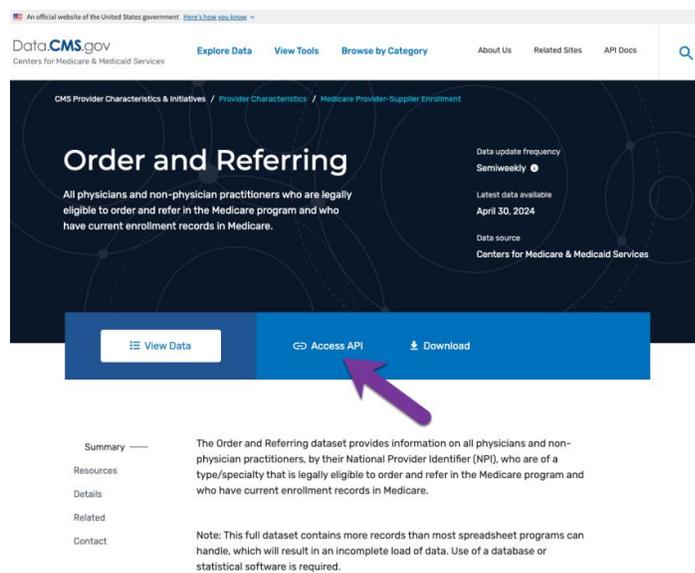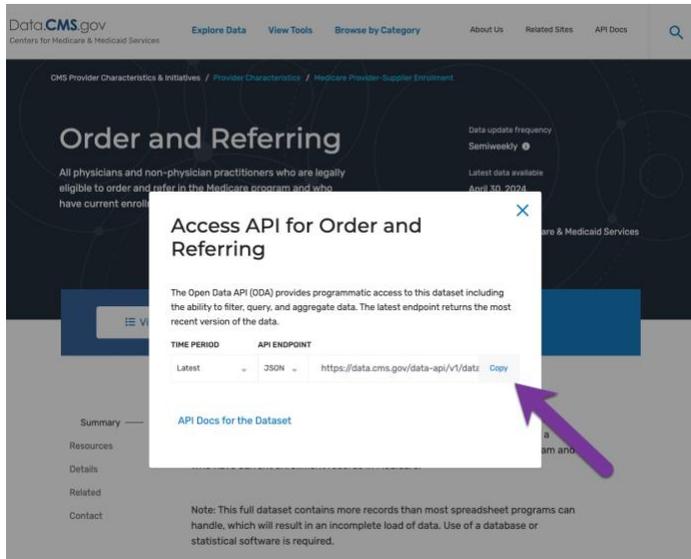
How do I get a link to the API endpoint from a dataset page?

If data is available via API, you will see an "Access API" button at the top of the dataset page. Clicking on the button will launch a pop-up with a link to the API endpoint.

For example, on this page you can click the button and copy the link.

Data.**CMS**.gov
Centers for Medicare & Medicaid Services

Using the API

What formats are available for the data?

Data is available on the API in JSON and JSON:API formats. You can also download the source CSV file directly.

I only get 1,000 rows, how do I get the entire dataset?

The default is to provide the first 1,000 rows per request. However, there is an ability to increase the limit to 5,000 rows per request.

You can use pagination to retrieve the entire dataset. For example, with our Opt Out Affidavits dataset:

1. Start with this request to get the number of rows https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data/stats
2. Make a request for the first 5,000 rows https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data?size=5000&offset=0
3. Make a request for the next 5,000 rows https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data?size=5000&offset=5000
4. Continue increasing the offset until you reach the count from step 1

Pagination Code Example

```
"""
Example in Python
Paginates through the Opt Out Affidavits data
https://data.cms.gov/provider-characteristics/medicare-provider-supplier-
enrollment/opt-out-affidavits
"""
import requests


url = "https://data.cms.gov/data.json"
title= "Opt Out Affidavits"
```

```
response = requests.request("GET", url)
response = response.json()
dataset = response['dataset']
for set in dataset:
    if title == set['title']:
        for distro in set['distribution']:
            if 'format' in distro.keys() and 'description' in distro.keys():
                if distro['format'] == "API" and distro['description'] ==
"latest":
                    latest_distro = distro['accessURL']

stats_endpoint = latest_distro + "/stats"

stats_response = requests.request("GET", stats_endpoint)
stats_response = stats_response.json()
total_rows = stats_response['total_rows']

i = 0
while i < total_rows:
    size = 5000
    offset_url = f"{latest_distro}?size={size}&offset={i}"

    offset = i
    offset_response = requests.request("GET", offset_url)
    print(f"Made request for {size} results at offset {i}")
    #if you want to save the data you would do that here
    i = i+size
```

How can I find the number of rows in a dataset?

You can use the data-viewer/stats or data/stats endpoint to accomplish this, such as
https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data-viewer/stats
or https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data/stats.

How do I download the data automatically?

You can use the download URL in the data.json to automatically download a dataset.

1. Start with a request to the data.json. Always start with the catalog, do not hard code or try to predict the download URL. The catalog will update automatically when new data is available.
2. Find the dataset within the catalog by matching the "*title"* with the name of the dataset.
3. Within each dataset, there is an array called *distribution* which will contain all the versions of the data in all available formats for each version.
4. The most recently release will be at the top of the *distribution* list, there is also a "*temporal"* field which can be used to find earlier releases.
5. Datasets are available as either CSV or ZIP files; this will be designated in the "*mediaType*" field. Please review the *distribution* list to determine what formats are available for the dataset.
6. The "*downloadURL"* will provide a direct download link for the data.

Data.**CMS**.gov
Centers for Medicare & Medicaid Services

CSV Download Link Code Example

```python
"""
Example in Python
Finds the download link for the latest Order and Referring Data
https://data.cms.gov/provider-characteristics/medicare-provider-supplier-
enrollment/order-and-referring
"""
import requests


url = "https://data.cms.gov/data.json"
title= "Order and Referring"
csv_distros =[]

response = requests.request("GET", url)
if response.ok:
    response = response.json()
    dataset = response['dataset']
    for set in dataset:
        if title == set['title']:
            for distro in set['distribution']:
                if 'mediaType' in distro.keys():
                    if distro['mediaType'] == "text/csv":
                        csv_distros.append(distro)
            latest = csv_distros[0]
            latest_url = latest['downloadURL']
            print(f"The latest CSV download link for {title} is
{latest_url}")
```

ZIP Download Link Code Example

```python
"""
Example in Python
Monthly Prescription Drug Plan Formulary and Pharmacy Network Information
https://data.cms.gov/provider-summary-by-type-of-service/medicare-part-d-
prescribers/monthly-prescription-drug-plan-formulary-and-pharmacy-network-
information
"""
import requests


url = "https://data.cms.gov/data.json"
title= "Monthly Prescription Drug Plan Formulary and Pharmacy Network
Information"
csv_distros =[]

response = requests.request("GET", url)
if response.ok:
    response = response.json()
    dataset = response['dataset']
    for set in dataset:
        if title == set['title']:
            for distro in set['distribution']:
                if 'mediaType' in distro.keys():
                    if distro['mediaType'] == "application/zip":
```

```
                csv_distros.append(distro)
        latest = csv_distros[0]
        latest_url = latest['downloadURL']
        print(f"The latest download link for {title} is {latest_url}")
```

What is the difference between the /data and /data-viewer endpoints?

Both endpoints will return the same data in different formats. The /data endpoint will return the data as a standard JSON while the /data-viewer endpoint will return data in a JSON:API format, which has a different response structure that includes metadata about the dataset.

Can I join across datasets?

You cannot join directly in the API. You can download the data independently or combine requests together to accomplish this, however, be aware that data column names and standards may vary.

Where can I find more information about the field names in the data?

The dataset page will contain a data dictionary with more information on the terms.

Can I make GRAPHQL or SOAP requests to the API?

Currently the API does not support those request types.

Filtering and Manipulating Data

Do you have examples of filters?

Equals

This example is an equals filter searching the Accountable Care Organizations 2021 dataset for a single ID.

https://data.cms.gov/data-api/v1/dataset/0ece71c6-1112-47b3-8f7a-5385087e7041/data?filter[filter-0-0][condition][path]=aco_id&filter[filter-0-0][condition][operator]=%3D&filter[filter-0-0][condition][value]=A4807

An equals filter can be simplified like this.

https://data.cms.gov/data-api/v1/dataset/0ece71c6-1112-47b3-8f7a-5385087e7041/data?filter[aco_id]=A4807

Keyword

The keyword search will look for matching words in every column. This example will check for "Alex" in the Order and Referring dataset. Notice that it finds matches on both the first and last name fields.

https://data.cms.gov/data-api/v1/dataset/c99b5865-1119-4436-bb80-c5af2773ea1f/data?keyword=alex

Multiple Conditions at Once

This search returns results from the Medicare Fee-For-Service Public Provider Enrollment dataset where the provider specialty is "PRACTITIONER - OPTOMETRY" and the location is Virginia.

https://data.cms.gov/data-api/v1/dataset/2457ea29-fc82-48b0-86ec-3b0755de7515/data?&filter[root-group][group][conjunction]=AND&filter[group-0][group][conjunction]=AND&filter[group-0][group][memberOf]=root-group&filter[filter-0-0][condition][path]=PROVIDER_TYPE_DESC&filter[filter-0-0][condition][operator]=%3D&filter[filter-0-0][condition][value]=PRACTITIONER%20-%20OPTOMETRY&filter[filter-0-0][condition][memberOf]=group-0&filter[filter-0-1][condition][path]=STATE_CD&filter[filter-0-1][condition][operator]=%3D&filter[filter-0-1][condition][value]=VA&filter[filter-0-1][condition][memberOf]=group-0

In

This search returns results from the Opioid Treatment Program Providers dataset where the provider is located from MD, MI, or VA with the results sorted by NPI.

https://data.cms.gov/data-api/v1/dataset/f1a8c197-b53d-4c24-9770-aea5d5a97dfb/data?filter[condition][path]=STATE&filter[condition][operator]=IN&filter[condition][value][]=MI&filter[condition][value][]=VA&filter[condition][value][]=MD&sort=NPI

Where can I find detailed documentation on filters?
You can find detailed documentation on building filter conditions here.

How can I sort my results?
You can use the sort query parameter to specify which column the results should be sorted by. These example show how to sort the Opt Out Affidavits dataset by NPI.

Lowest first https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data?sort=NPI

Highest first https://data.cms.gov/data-api/v1/dataset/9887a515-7552-4693-bf58-735c77af46d7/data?sort=-NPI

I only want some of the columns, how do I do that?
You can use the column query parameter to limit the columns that are returned to you by providing a comma separated list of the columns to include.

This example will return just the NPI, First Name, and Last Name from the Order and Referring dataset.

https://data.cms.gov/data-api/v1/dataset/c99b5865-1119-4436-bb80-c5af2773ea1f/data?column=NPI,FIRST_NAME,LAST_NAME

How do I specify the number of results I want?
You can use the size and offset parameters to accomplish this.

The size parameter specifies the number of results to return back (up to 5,000 at a time) and the offset parameter specifies a starting point.

For example, this query will return results 501-600 https://data.cms.gov/data-api/v1/dataset/c99b5865-1119-4436-bb80-c5af2773ea1f/data?size=100&offset=500.

What filter conditions are available?

Our API support the following:

- =
- <>
- >
- >=
- <
- <=
- STARTS_WITH
- CONTAINS
- ENDS_WITH
- IN
- NOT IN
- BETWEEN
- NOT BETWEEN
- IS NULL
- IS NOT NULL

Troubleshooting

Why does my request that worked last week fail now?

You may be using a version of the data that is no longer available. Refer to the question on "Is there a way to always get the latest data."

Why isn't this dataset available via the API?

Some datasets are non-interactive and can only be downloaded, either from the dataset page or from the data.json.

Who can I contact if I have a question about the data or found an error in the data?

Please use the "Contact Email" on the dataset's page to contact the data owner.

Who can I contact if I have questions about using the API?

Please contact data.support@cms.hhs.gov for technical questions.

Why am I getting an access denied message?

Please contact data.support@cms.hhs.gov for assistance and include the reference number from the error message in your email.

When will the data be updated next?

The dataset landing page will include information on the update interval and when the data was last published. If you have further questions on the timing of the next release, please contact the data owners using the "Contact Email" on the dataset's page.

Why do some datasets have multiple versions but others do not?

This is at the discretion of the dataset owner. Please contact the data owner using the "Contact Email" on the dataset page if you have feedback. Some datasets include the historical data and the most recent data in single version, so historical data may still be available to you.

Why are fields different between versions?

Data can differ between versions for a number of reasons, such as different standards, identifying new data to be collected, identifying existing data as no longer needed, and different collection methods. This is at the discretion of the data owner. Please contact the data owner using the "Contact Email" on the dataset page if you have feedback.

Why is all data represented as a string?

This is to ensure backwards compatibility with legacy systems.

Can you provide me with code to use the API?

No, we can provide explanations on how to access the data but we do not provide development support.

I have a long request that is not working?

We have a limit of 8,192 characters in the URL for your request. If you encounter this limit please contact data.support@cms.hhs.gov for assistance with simplifying your query.

How do I search for fields with spaces in the name?

You can replace spaces with %20. Additional documentation on other encodings can be found here.

For example this filter for "PROVIDER TYPE TEXT - BUYER" equaling "PART A PROVIDER - HOSPITAL" for the Hospital Change of Ownership dataset could be accomplished as https://data.cms.gov/data-api/v1/dataset/c04031db-54ce-461c-85d1-d2613d71f167/data?&filter[PROVIDER%20TYPE%20TEXT%20-%20BUYER]=PART%20A%20PROVIDER%20-%20HOSPITAL

How can I find answers to general site related questions?

The "About Us" page provides general information on using the site.