

Install a Provider and add a Connection

Apache Airflow is built in modular way. The “core” of Apache Airflow provides core scheduler functionality which allow you to write some basic tasks, but the capabilities of Apache Airflow can be extended by installing additional packages, called providers. Providers can contain operators, hooks, sensor, and transfer operators to communicate with a multitude of external systems, but they can also extend Airflow core with new capabilities.

Prerequisites

To add a provider you need have set up your local Astro project as described under the section “Set up your local environment”

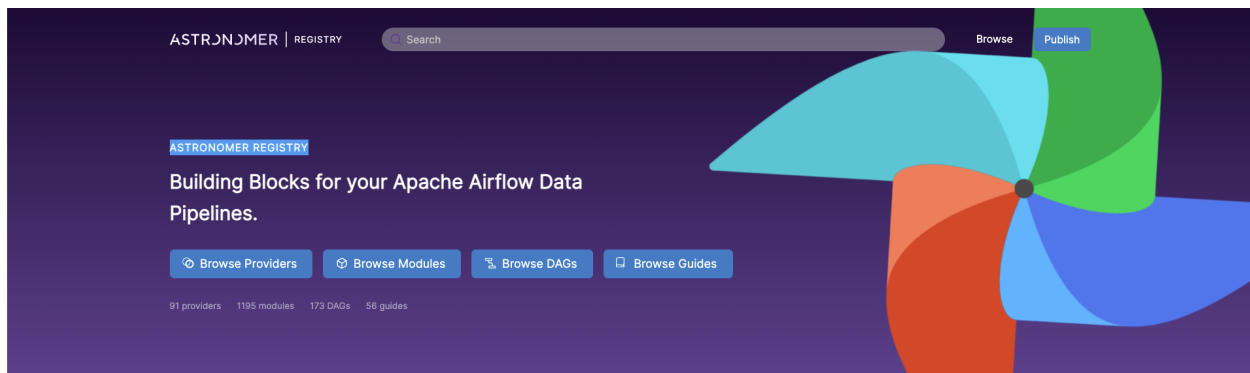
Where to find a Provider?

To interact with an external system such as AWS, DBT or Snowflake, you have to install the corresponding Provider. Airflow has 90+ providers.

The easiest way to find the one you are looking for, go to the <https://registry.astronomer.io/>

The Astronomer Registry is the discovery and distribution hub for Apache Airflow integrations.

You will find data pipeline examples, guides, documents and more.



Providers

Python packages containing all relevant Airflow modules for a third-party service.

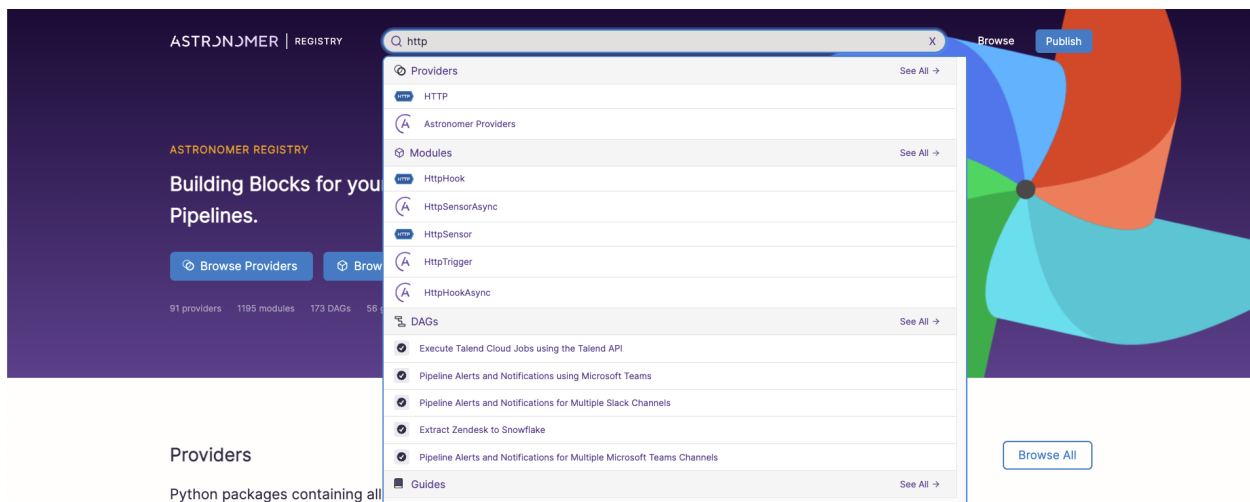
[Browse All](#)

Find the HTTP provider

To interact with an API, a website or any HTTP resource, you have to install the http provider.

Go to <https://registry.astronomer.io/>

Look for **http** in the search bar



Select **HTTP** under **Providers**

The screenshot shows the Apache Airflow Providers Registry page for the HTTP provider. The header is dark purple with 'ASTRONOMER | REGISTRY' and a search bar. Below the header, there's a navigation bar with 'Providers / HTTP' and tabs for 'Overview', 'Docs', 'Example DAGs', and 'Connections'. The main content area has a 'HTTP' badge and tabs for 'Core' and 'Web Service'. It describes the provider as 'An Apache Airflow provider for HTTP.' Below this, there are four boxes: 'VERSION 4.0.0', 'DOWNLOADS 2,222,717/month', 'LAST PUBLISHED Jul. 16, 2022', and 'Quick install' with a code box containing 'pip install apache-airflow-providers-http==4.0.0'. At the bottom, there's an 'Available Modules' section with a description and tabs for 'Hooks', 'Operators', and 'Sensors'. On the right, there are 'Helpful Links' including 'View on GitHub', 'Using Providers', and 'Learn Airflow'.

On this page, you get access to Provider's informations such as:

- Version
- Popularity
- Last update
- Quick install
- Available modules: ways of interacting with HTTP from your data pipelines.
- Example DAGs: data pipeline examples

Click on the bar under `Quick install` to copy the `pip install` line

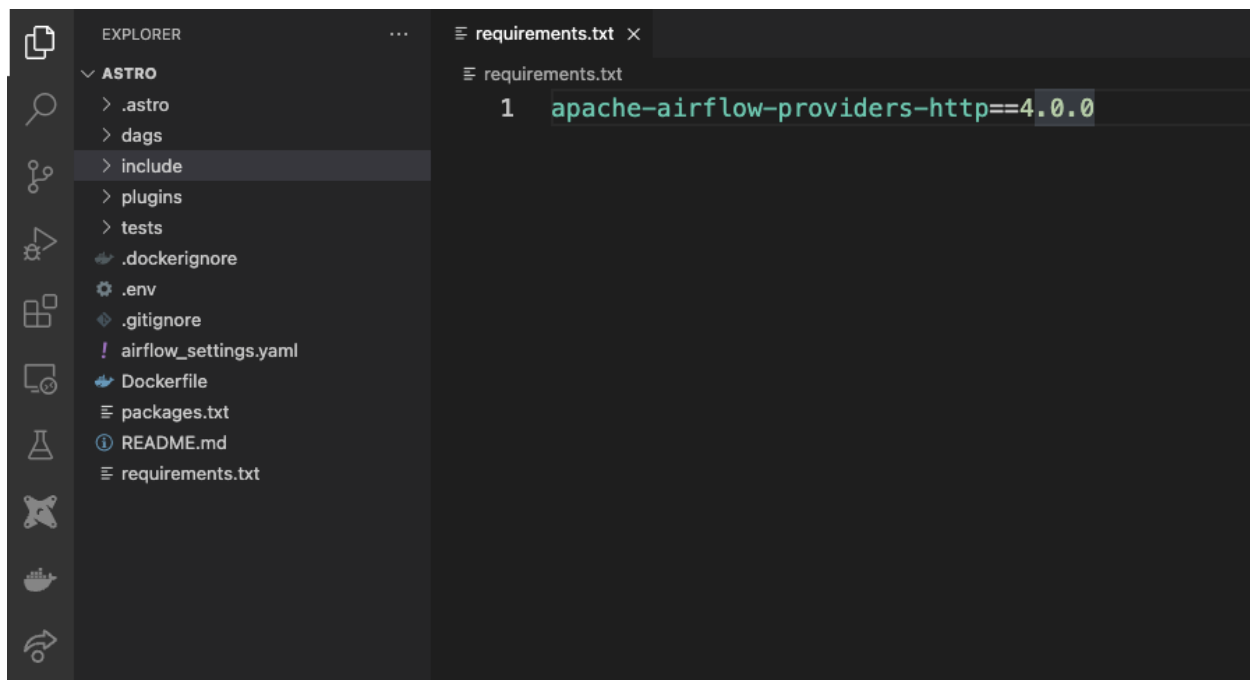
Install the HTTP provider

In your local Astro project

Open the file `requirements.txt`

Paste what you copied just before. Should be:

```
apache-airflow-providers-http==4.0.0
```



Save the file

Restart the Astro project by typing in your terminal: `astro dev restart` (make sure you are in the folder where the Astro project is)

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL
~/Documents/astro astro dev restart
Env file ".env" found. Loading...
[+] Running 4/4
:: Container astro_a04e39-triggerer-1 Stopped
:: Container astro_a04e39-webserver-1 Stopped
:: Container astro_a04e39-scheduler-1 Stopped
:: Container astro_a04e39-postgres-1 Stopped
[+] Building 9.9s (8/10)
=> [internal] load .dockerignore
=> => transferring context: 34B
=> [internal] load metadata for quay.io/astronomer/astro-runtime:5.0.6
=> [1/1] FROM quay.io/astronomer/astro-runtime:5.0.6@sha256:9d0981d94e00a6bc7c28b55df3276fd596a545d8cca1c
=> [internal] load build context
=> => transferring context: 1.92kB
```

Once your Airflow instance is up and running again, execute the command

```
astro dev run providers list
```

You should see the http provider

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL

| - Google Cloud (GCP) https://cloud.google.com/ |
| - Google Firebase https://firebase.google.com/ |
| - Google LevelDB https://github.com/google/leveldb/ |
| - Google Marketing Platform https://marketingplatform.google.com/ |
| - Google Workspace https://workspace.google.com/ (formerly Google Suite) |
| Hypertext Transfer Protocol (HTTP) https://www.w3.org/Protocols/ | 4.0.0
| Internet Message Access Protocol (IMAP) https://tools.ietf.org/html/rfc3501 | 3.0.0
| Microsoft Azure https://azure.microsoft.com/ | 4.0.0
| PostgreSQL https://www.postgresql.org/ | 5.0.0
| Redis https://redis.io/ | 3.0.0
| Snowflake https://www.snowflake.com/ | 3.0.0
| SQLite https://www.sqlite.org/ | 3.0.0

apache-airflow-providers-http
apache-airflow-providers-imap
apache-airflow-providers-microsoft-azure
apache-airflow-providers-postgres
apache-airflow-providers-redis
apache-airflow-providers-snowflake
apache-airflow-providers-sqlite
```

Well done! You have successfully installed a Provider to Airflow 😎

Add a Connection

Create a Connection in the UI

- Navigate in the Airflow UI to **Admin** then **Connections** , click the blue plus button
- Fill it in like the following (note **https** is the Schema due to legacy reasons)
 - connection_id: **github_api**
 - host: **api.github.com**
 - schema: **https**

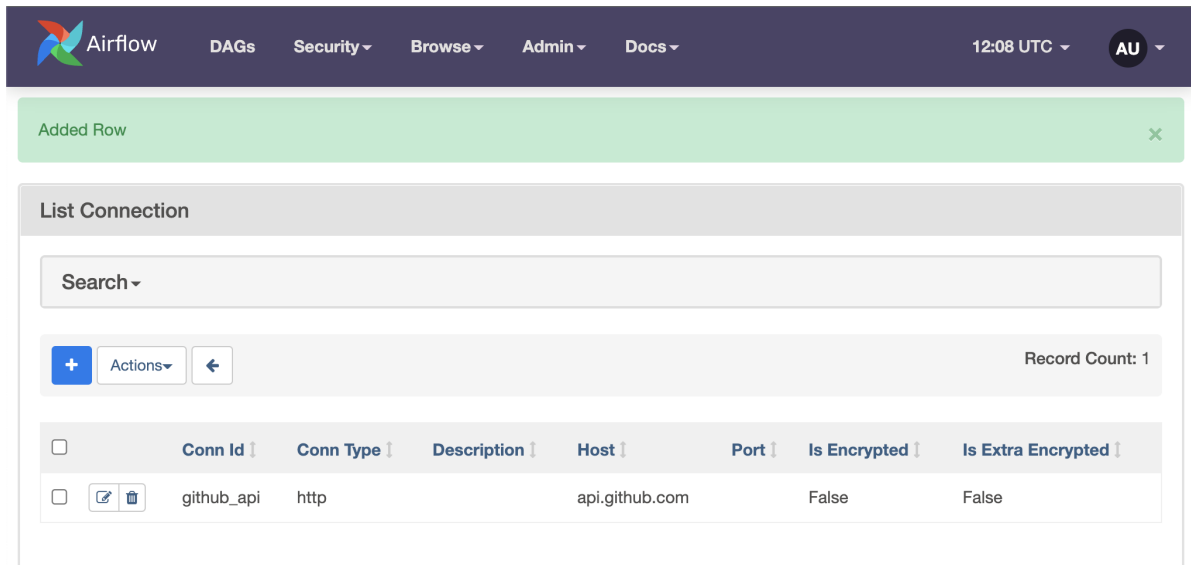
Add Connection

Connection Id *	github_api
Connection Type *	HTTP
Connection Type missing? Make sure you've installed the corresponding Airflow Provider Package.	
Description	
Host	api.github.com
Schema	https
Login	
Password	
Port	
Extra	

Save Test ↩

Click **Save**

You should see a new Connection as shown below:



The screenshot shows the Airflow web interface. At the top, there's a navigation bar with the Airflow logo, 'DAGs', 'Security', 'Browse', 'Admin', and 'Docs' menus. The time is 12:08 UTC and the user is 'AU'. Below the navigation bar, a green banner says 'Added Row'. The main content area is titled 'List Connection'. It has a search bar, a '+ Actions' button, and a 'Record Count: 1' indicator. Below this is a table with columns: Conn Id, Conn Type, Description, Host, Port, Is Encrypted, and Is Extra Encrypted. The table contains one row with the following data: Conn Id: github_api, Conn Type: http, Host: api.github.com, Port: (empty), Is Encrypted: False, Is Extra Encrypted: False.

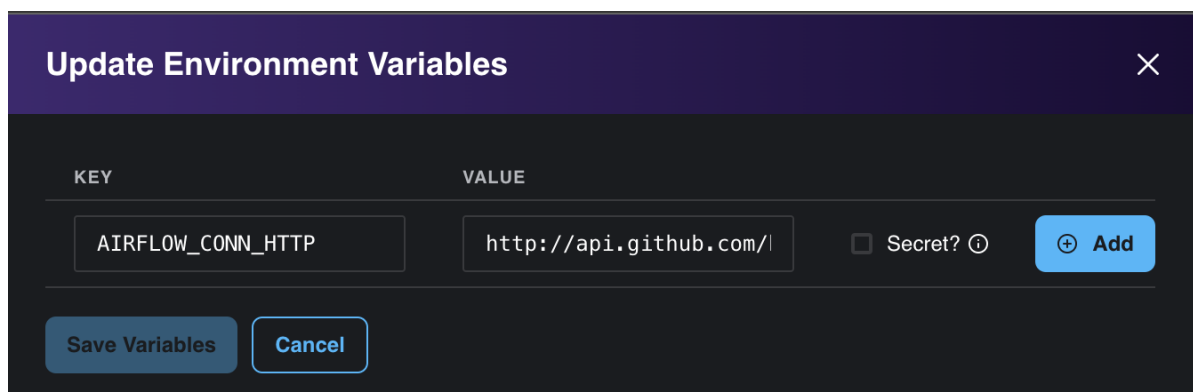
Conn Id	Conn Type	Description	Host	Port	Is Encrypted	Is Extra Encrypted
github_api	http		api.github.com		False	False

Other ways to create a connection

- Additionally, connections can be created via a text form in your `.env` file, structured as a URI or JSON, e.g. the URI form would be:

```
AIRFLOW_CONN_GITHUB_API="http://api.github.com/https"
```

- NOTE: The `.env` file only applies to your Local Airflow, but can be added to the Astronomer UI via the `Environmental Variables` section on your Deployment, which you can read more about [here](#). Make sure to Hit `Add` and `Save Variables` to add them to your Astronomer Cloud Deployments.



The screenshot shows a modal dialog titled 'Update Environment Variables'. It has a table with two columns: 'KEY' and 'VALUE'. The first row has the key 'AIRFLOW_CONN_HTTP' and the value 'http://api.github.com/'. To the right of the value is a checkbox labeled 'Secret?' with an information icon. At the bottom right of the table is a blue button with a plus icon and the text 'Add'. Below the table are two buttons: 'Save Variables' and 'Cancel'.

KEY	VALUE
AIRFLOW_CONN_HTTP	http://api.github.com/

- You can also add connections via a Secrets Backend and in your Dockerfile, but these are advanced topics and not covered here

Additional resources

Astronomer Registry: <https://registry.astronomer.io/>

Airflow providers: <https://airflow.apache.org/docs/apache-airflow-providers/>

Airflow Http Connection: [HTTP Connection](#) | [Airflow OSS Doc](#)