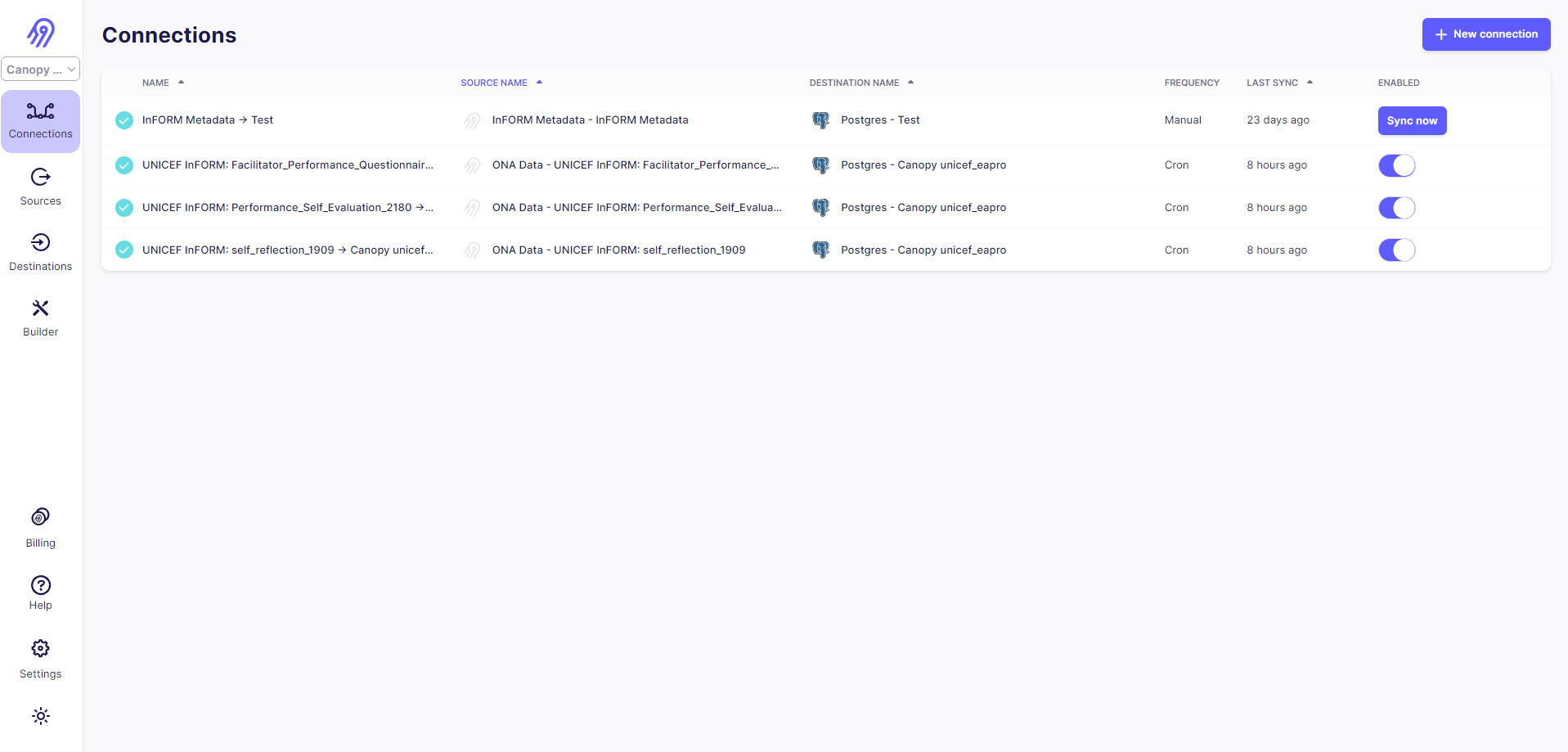
Airbyte Cloud Guide

Extracting and loading data to the database

Once data has been uploaded to the forms on InForm, the data has to be loaded into a database for further transformation to be done. Ona uses Airbyte to load data into a database. [Airbyte](https://airbyte.com/) is an open-source data integration platform that securely extracts data from different data sources and reliably loads the data to a data warehouse or database. Clients only have access to Airbyte if they have access to their own deployment of Airbyte. Ona does not currently provide access to their Airbyte instance. This document will only be relevant for clients with their own instances of Airbyte.

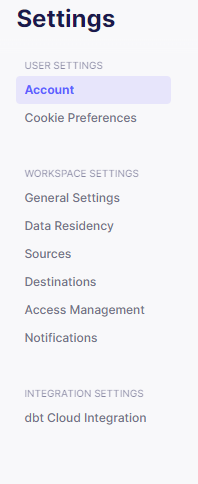
## Getting started with Airbyte Cloud

To create the connectors on Airbyte, an invitation to the Canopy UNICEF Inform workspace on Airbyte Cloud will be sent via email to the relevant parties. After receiving an invitation, click on the verification link, and you will be redirected to the Canopy UNICEF InForm workspace.



Within the workspace, there are different sections:

1. **Connections:** This section contains the created data pipelines. A connection is an automated data pipeline that replicates data from a source to a destination, allowing users to specify the replication frequency.
2. **Sources:** This section has the defined sources. A source can be an API, a database, or a file that you want to ingest your data from. For this project, InForm is our source.
3. **Destinations**: A destination is a database, data warehouse, or analytics tool into which you want to load your extracted data pulled from a source.
4. **Billing:** This section shows the total credits usage, usage per connection, etc.
5. **Settings*:***Within this section, we have different sub-sections:



1. **Account:** Displays the name and email address associated with the account. One can also update the account password from this section.
2. **Cookie preferences:** Allows you to set what you want the retrieved browser data to be used for.
3. **General settings:** Displays the workspace name.
4. **Data residency:** This allows you to choose the default data processing location for all your connectors.
5. **Sources:** Displays the available source connectors.
6. **Destinations:** Displays the available destination connectors.
7. **Access management:** Displays the users who have access to the workspace.
8. **Notifications:** Allows you to set up the actions that you would like to receive notifications for such as monitoring sync failures.

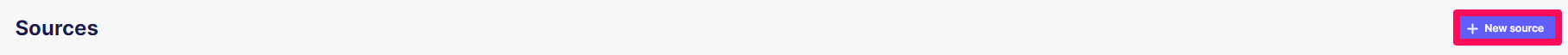
### Setting up an OnaData source

To set up a source, see the step-by-step guide below:

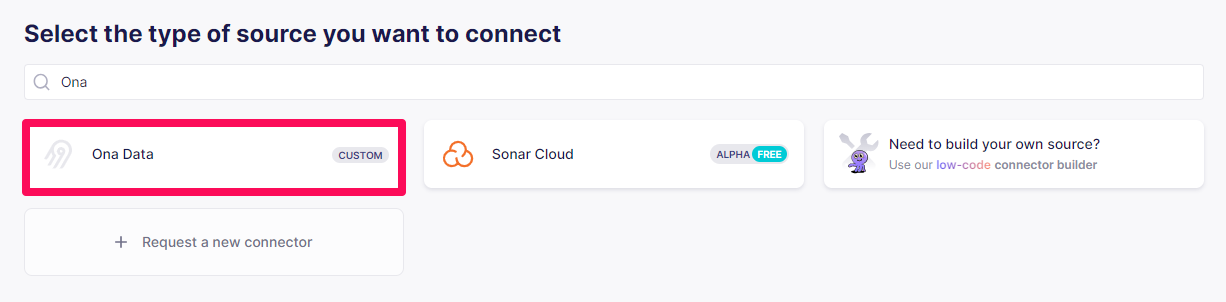
1. Click on **Sources**.



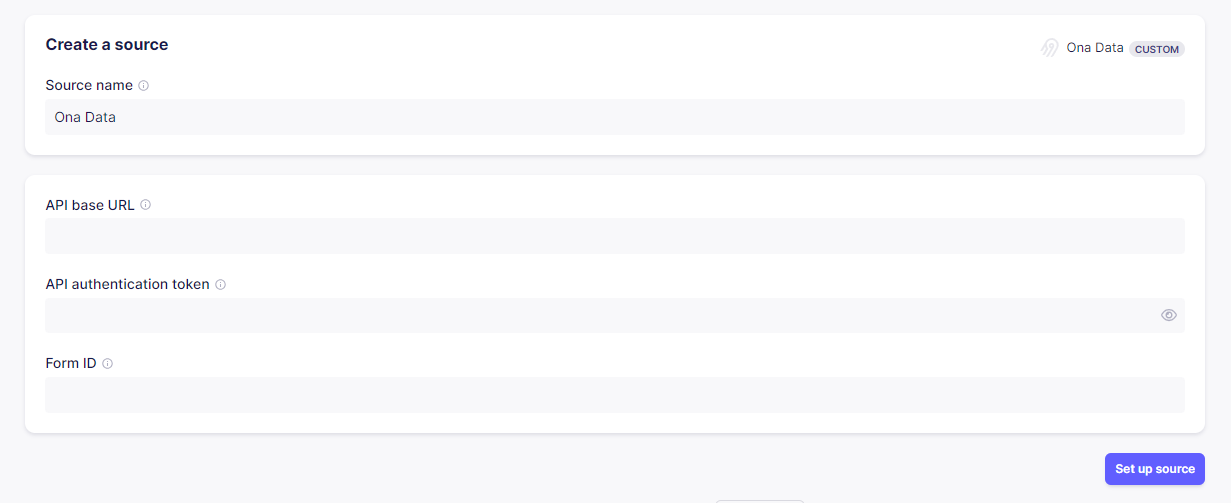
1. On the **Sources** page, click on **+New Source**.



1. On the **Set up the source** page, select the source you want to set up from the source type dropdown. To ingest data from InForm, select **Ona Data** under the sources available.



1. On the **Create a source** page, specify the following key components:



1. **Source name:** This could be the name of the form, such as Social Mobilization Indicators.
2. **API base URL:** This is the API URL. The InForm URL is <https://data.inform.unicef.org>.
3. **API authentication token:** This token is retrieved from the InForm account settings page. For more information on accessing your API token, see this [help documentation](https://help.ona.io/knowledge-base/connect-a-form-to-powerbi-using-the-api/#get-api-key).



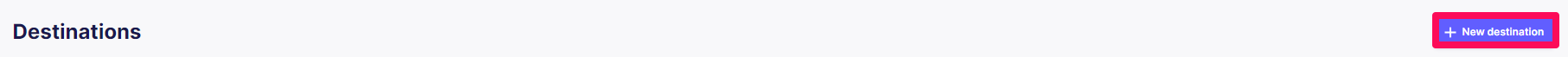
1. **Form ID:** The form ID is retrieved from the form URL on InForm. After navigating to the form's overview page, the form ID is the last part of the form’s overview page URL. For example, if the form’s URL is <https://inform.unicef.org/uniceftemplates/635/765>, 765 is the form id.

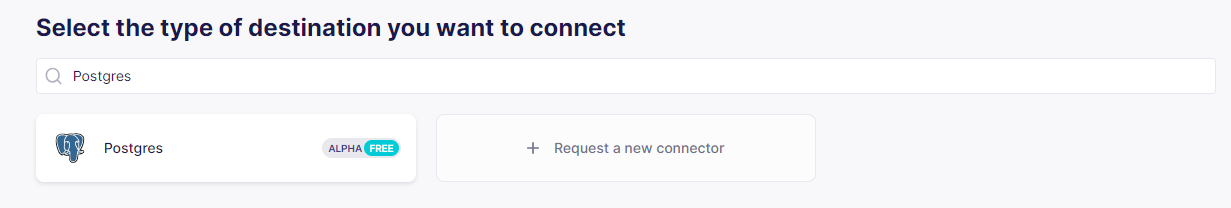
5. Once you have submitted the details above, click on **Set up source** to create the source.

### Setting up a Postgres destination

For this InForm workaid, the destination is a Postgres database. However, one can set up different destinations available on Airbyte. In this section, we will look at how to set up a Postgres destination:

1. On the **Destinations** page, click on **+ New destination**.

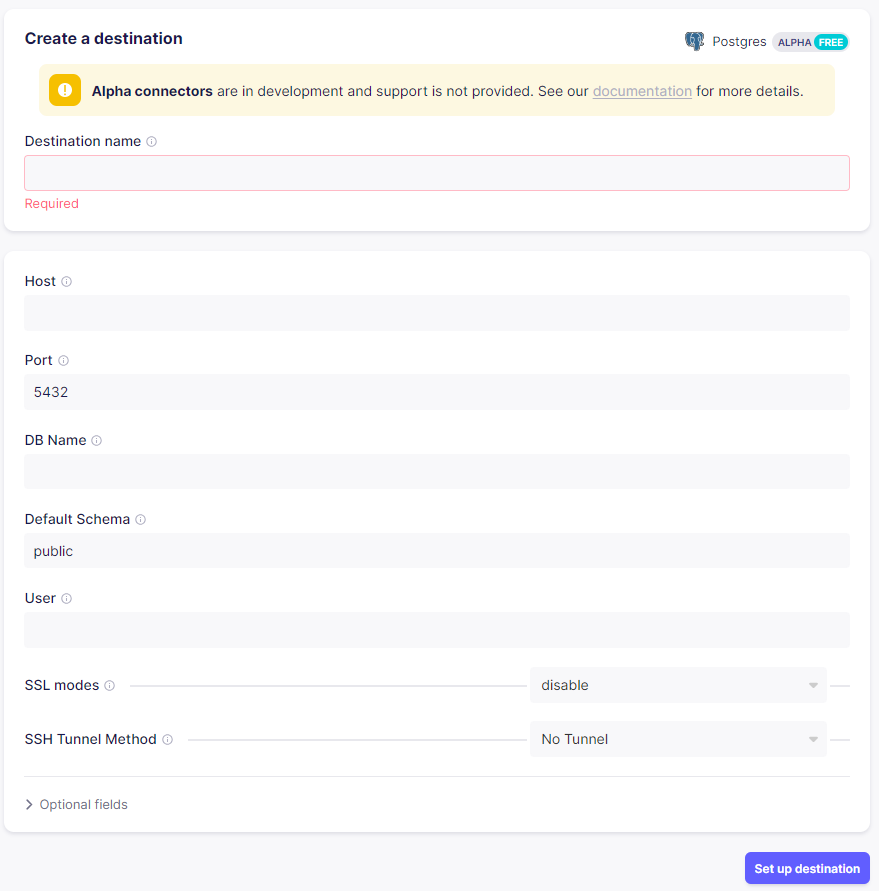


1. On the **Setup destination** page, select **Postgres**.

The Canopy UNICEF InForm workspace has the Postgres database destination set up. When using this workspace, you do not need to specify a new database if the goal is to have the data within the InForm database. If you are using another database, you will need to create a new database.

#### Set up a new database

To set up a new database, follow the step-by-step guide:

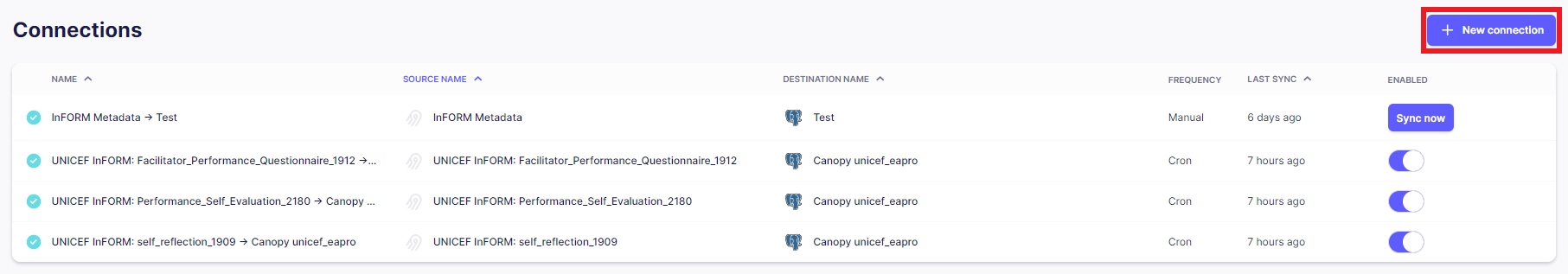
1. Specify the database details under the *Create a destination* page. For a Postgres destination, specify the following:
2. **Destination name:** Name of the destination. For instance, the database name.
3. **Host:** Hostname of the database.
4. **Port:** Port of the database.
5. **DB Name:** Name of the database.
6. **Default Schema:** The default schema the tables are written to.
7. **User:** Username to access the database.
8. **SSL modes:** SSL connection modes. We highly recommend selecting **Require** or at least **Prefer**.
9. **SSH Tunnel Method:** Whether to initiate an SSH tunnel as the authentication method.
10. Then, click on **Set up destination** to create the destination. 

### Setting up a connection

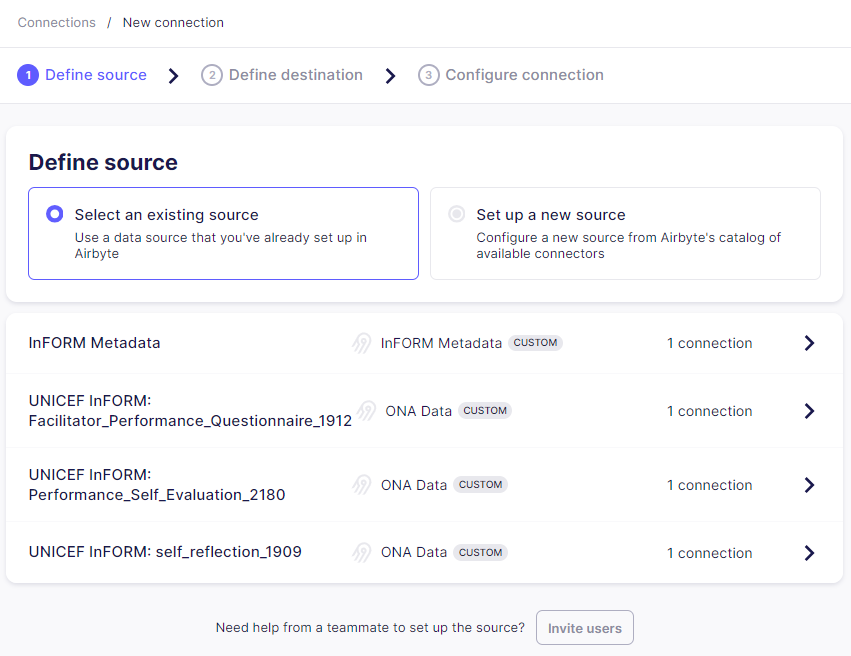
A connection is an automated data pipeline that replicates data from a source to a destination. You can create a connection from an existing source or destination or set up a new source or destination while creating the connection.

To set up a connection, see the following steps below:

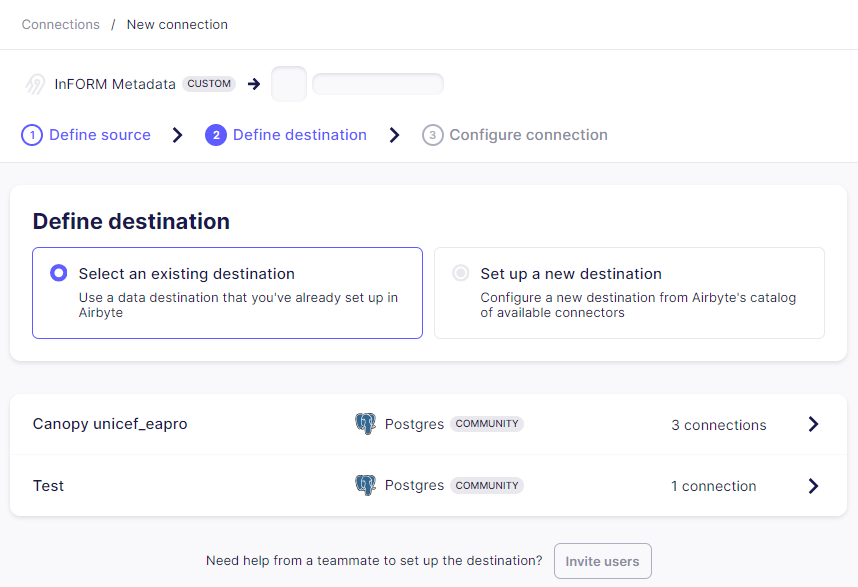
1. Click on **+ New connection** under the connections page.



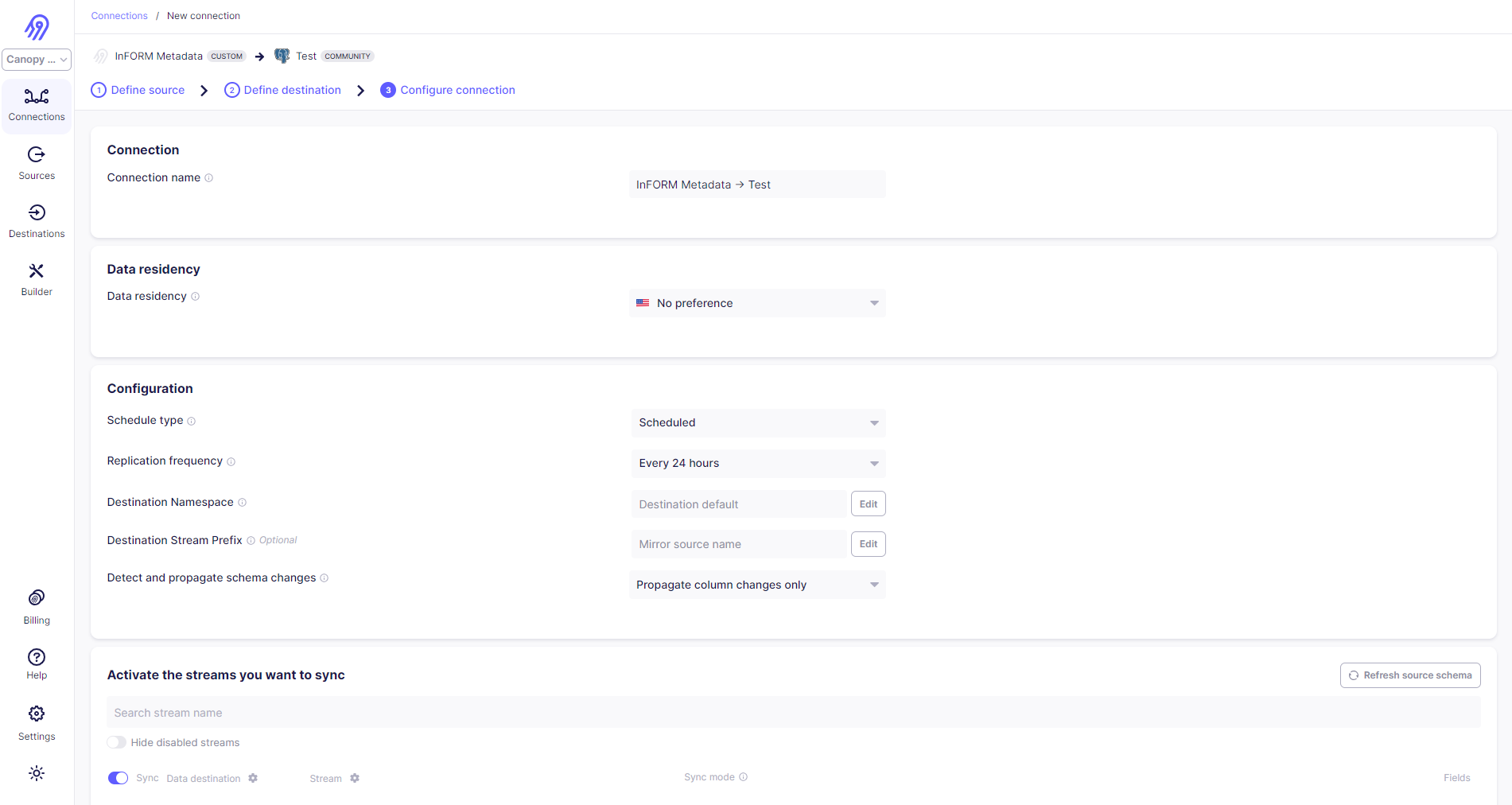
1. On the define source page, select an existing source or set up a new source. To set up a new source, take a look at [setting up a source.](#_pbs3erpskkcw)



1. Under the define destination page, select a destination. You can select an existing destination or set up a new destination. To set up a new destination, take a look at [setting up a destination](#_lsiczsic0zua).

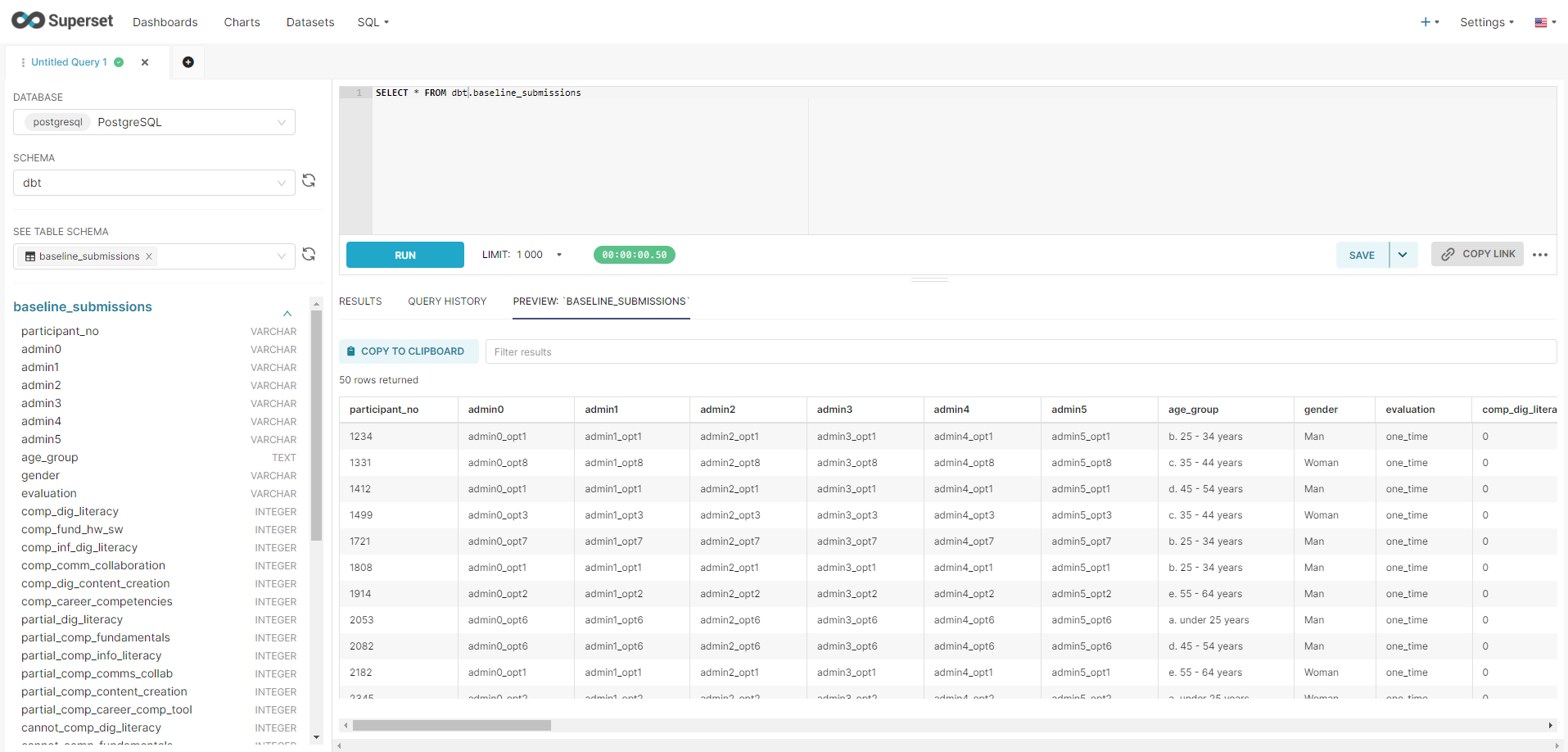


1. Once you select a destination, you will be redirected to the setup connection page.



1. From the **Replication frequency** dropdown, select how often you want the data to sync from the source to the destination. The default replication frequency is **Every 24 hours**.
2. From the **Destination Namespace**, select the destination location in which you want to store the data. The default configuration is **Destination default**.
3. In the **Destination Stream Prefix** (Optional) field, add a prefix to stream names. For instance, when you add *templates\_* as the prefix, the stream-labeled cases will be renamed to *templates\_cases* in the database.
4. Under the **Activate the streams you want to sync** section, configure the sync settings:
5. Toggle the **Sync** button to enable sync for the stream.
6. Sync mode: select how you want the data to be replicated from the source to the destination.
7. For the source, select **Full refresh** to copy the entire dataset each time you sync.
8. For the destination, select:
   1. **Overwrite** to erase the old data and replace it completely.
   2. **Append** to capture changes to your table.
      1. **Note**: This creates duplicate records.
9. Click **Set up connection**, and Airbyte tests the connection. If the sync is successful, the connection page is displayed.

Once the connectors have been successfully set up, you can access the tables within the Airbyte schema on the SQL Lab as shown below:



**Note:** Only users with access to the SQL Lab will be able to view and run queries against the database.

For more information on Airbyte Cloud, take a look at this [documentation](https://docs.airbyte.com/?_gl=1*1l72vob*_gcl_au*MTg2NTE3NzA4My4xNzAwNjM1OTc5).