

PostgreSQL Import Integration

[Learn more about PostgreSQL Export Integration.](#)

The Data Connector for PostgreSQL enables you to directly import data from your PostgreSQL to Treasure Data.

For sample workflows showing how to import data from your PostgreSQL, view [Treasure Boxes](#).

Continue to the following topics:

- [Prerequisites](#)
- [Using TD Console](#)
- [Transfer Your PostgreSQL Account Data to Treasure Data](#)

Prerequisites

- Basic knowledge of Treasure Data
- Basic knowledge of PostgreSQL
- A PostgreSQL instance running **remotely**, for example on RDS.

Using TD Console

Create a New Connection

When you configure a data connection, you provide authentication to access the integration. In Treasure Data, you configure the authentication and then specify the source information.

1. Open **TD Console**.
2. Navigate to **Integrations Hub > Catalog**
3. Search for and select PostgreSQL. Select **Create**.



4. The following dialog opens.

New Authentication
PostgreSQL
✕

1 Credentials >
 2 Details

Host:

Port:

User:

Password:

Use SSL:

▼ OPTIONS

These options are passed directly to the JDBC connection driver.

Add

Socket connection timeout:
Timeout (in seconds) for socket connect. 0 means no timeout.

Network timeout:
Timeout (in seconds) for network socket operations. 0 means no timeout.

[Learn more](#)
Continue

5. Enter the required credentials and set the parameters. Select **Continue**.

Host	The host information of the source database, such as an IP address.
Port	The connection port on the source instance. The PostgreSQL default is 5432.
User	Username to connect to the source database.
Password	The password to connect to the source database.
Use SSL	Check this box to connect using SSL
Specify SSL version	Select which SSL version to use for the connection.
Socket connection timeout	Timeout (in seconds) for socket connection (default is 300).
Network timeout	Timeout (in seconds) for network socket operations. 0 means no timeout.

Name the Connection

1. Type a name for your connection. If you would like to share this connection with other users in your organization, select **Share with others**. If this box is unchecked, then the connection is visible only to you.
2. Select **Done**.

Transfer Your PostgreSQL Account Data to Treasure Data

After creating the authenticated connection, you are automatically taken to Authentications.

1. Search for the connection you created.
2. Select **New Source**. The Create Source dialog opens.

Connection

1. Type a name for your **Source** in the Data Transfer field.

1 Connection	Data Transfer Name:	<input type="text"/>
2 Source Table	Authentication:	01_se_postgres_db ▼
3 Data Settings		
4 Data Preview		
5 Data Placement		

2. Click **Next**.

Source Table

1. Edit the following parameters

1 Connection	Database name:	lmbtwnww
2 Source Table	Schema:	public
3 Data Settings	Use custom SELECT query?:	<input type="checkbox"/> Use if you need more than a simple SELECT (columns) FROM table WHERE (condition).
4 Data Preview	SELECT columns:	* comma-separated list of columns to SELECT
5 Data Placement	Table:	saturday
	WHERE condition:	<input type="text"/>
	ORDER BY:	<input type="text"/>

Parameters	Description
Database name	The name of the database you are transferring data from. For example, your_database_name.
Use custom SELECT query?	Use if you need more than a simple SELECT (columns) FROM table WHERE (condition).
SELECT columns	If there are only specific columns you would like to pull data from, list them here. Otherwise, all columns are transferred.
Table	The table from which you want to import the data.
WHERE condition	If you need additional specificity on the data retrieved from the table you can specify it here as part of WHERE clause.
ORDER BY	Specify if you need the records ordered by a particular field.

Data Settings

1. Select **Next**. The Data Settings page opens.
2. Optionally, edit the data settings or skip this page of the dialog.

1 Connection

2 Source Table

3 Data Settings

4 Data Preview

5 Data Placement

Optionally, you can modify data settings and then see your changes in Data Preview. [Skip This Step](#)

Incremental?:
When run repeatedly, attempt to only import new data since the last import

Rows per batch:
Number of rows to fetch one time (used for java.sql.Statement#setFetchSize)

Default timezone:

After SELECT:
This SQL will be executed after the SELECT query in the same transaction.

▼ COLUMN OPTIONS

Advanced options for individual columns.

[Add](#)

Parameters	Description
Incremental:	When you want to repeatedly run this transfer, select this checkbox to import data only since the last time the import was run.
Rows per batch	Extremely large datasets can lead to memory issues and subsequently failed jobs. Use this flag to breakdown the import job into batches by the number of rows to reduce the chances of memory issues and failed jobs.
Default timezone	The timezone to be used when doing the import.
After SELECT	This SQL is executed after the SELECT query in the same transaction.
Column Options	Select this option to modify the type of column before importing it. Select Save to save any data setting you have entered.
Default Column Options	Select this option to define the data type according to default SQL types before importing it. Select Save to save any data settings you have entered. This option is not available in the TD Console. Set this option using TD CLI or TD Workflow

Preview

You can see a [preview](#) of your data before running the import by selecting Generate Preview.

Data shown in the data preview is approximated from your source. It is not the actual data that is imported.

1. Click **Next**.
Data preview is optional and you can safely skip to the next page of the dialog if you want.
2. To preview your data, select **Generate Preview**. Optionally, click **Next**.
3. Verify that the data looks approximately like you expect it to.

Create Source
Using onetrust_demo

The preview shows a subset of data from the source based on the data settings. Refer to [help document](#) to learn more about preview data.

8 columns

	Ab_id	Ab_language	Ab_identifier	last_updated_date	Ab_link_token	
1	f7abf910-b5da-47c2-bbee-37f4c86...	NULL	Quan3	2020-09-25 22:42:59...	NULL	0
2	9022117f-cf3c-418c-b527-a8bd9a9...	NULL	Quan2	2020-08-05 03:48:19...	NULL	0
3	a432b52f-3d93-483b-b65f-3c7530...	NULL	Quan4	2020-08-05 03:48:19...	NULL	0
4	233ec0c2-70ab-4de4-ac48-a4a048f...	NULL	Quan5	2020-08-05 03:48:19...	NULL	0
5	f78be70b-8b5d-404e-b663-b606a2...	NULL	Quan1	2020-08-05 03:48:19...	NULL	0
6	db5d9f89-c264-4d82-a246-5939e5...	NULL	example@otprivacy.com	2020-08-06 17:51:12...	NULL	0
7	5ef9542c-315d-4b56-ad1c-c63ad0...	NULL	Michael.White@gmail.com	2020-09-09 20:01:45...	NULL	0
8	3f1dfcb9-1904-4517-9087-0cc45f0...	NULL	Robert.Brown@gmail.com	2020-09-09 20:01:45...	NULL	0
9	4a3a88dd-11a3-4c8b-a1d9-d7043f...	NULL	Mary.Anderson@gmail.com	2020-09-09 20:01:46...	NULL	0
10	4f69893a-9e49-46dc-9519-1cf9dea...	NULL	Elizabeth.Scott@gmail.com	2020-09-09 20:01:47...	NULL	0
11	33342e5d-4c95-4cfe-a622-4e91dc5...	NULL	David.Miller@aol.com	2020-09-09 20:01:47...	NULL	0
12	f54b0d7c-df75-4bf3-934a-dc19a96...	NULL	Robert.Anderson@att.com	2020-09-10 04:57:16...	NULL	0
13	43bfe156-dfba-43b8-964d-1b2a4ae...	NULL	Elizabeth.Miller@google.com	2020-09-10 04:57:16...	NULL	0

Cancel Back Next

4. Select **Next**.

Data Placement

For data placement, select the target database and table where you want your data placed and indicate how often the import should run.

1. Select **Next**. Under **Storage** you will create a new or select an existing database and create a new or select an existing table for where you want to place the imported data.

1 Connection

2 Source Table

3 Data Settings

4 Data Preview

5 **Data Placement**

▼ STORAGE

Database: chung_default_db

Table: sftp_v2_devproxy

Method:

- Append: Add records into existing table.
- Always Replace: Always clear the destination table before adding records.
- Replace on new data: When there is new data, delete existing data, and insert new data.

Timestamp-based Partition Key: time

Data Storage Timezone: UTC (default)

▼ SCHEDULE

Repeat: Off On

Scheduling Timezone: Asia/Saigon

2. Select a **Database** > **Select an existing** or **Create New Database**.
3. Optionally, type a database name.
4. Select a **Table** > **Select an existing** or **Create New Table**.
5. Optionally, type a table name.
6. Choose the method for importing the data.
 - **Append** (default)-Data import results are appended to the table. If the table does not exist, it will be created.
 - **Always Replace**-Replaces the entire content of an existing table with the result output of the query. If the table does not exist, a new table is created.
 - **Replace on New Data**-Only replace the entire content of an existing table with the result output when there is new data.
7. Select the **Timestamp-based Partition Key** column.
If you want to set a different partition key seed than the default key, you can specify the long or timestamp column as the partitioning time. As a default time column, it uses upload_time with the add_time filter.

8. Select the **Timezone** for your data storage.
9. Under **Schedule**, you can choose when and how often you want to run this query.

- Run once:
 - a. Select **Off**.
 - b. Select **Scheduling Timezone**.
 - c. Select **Create & Run Now**.
- Repeat the query:
 - a. Select **On**.
 - b. Select the **Schedule**. The UI provides these four options: *@hourly*, *@daily* and *@monthly* or custom *cron*.
 - c. You can also select **Delay Transfer** and add a delay of execution time.
 - d. Select **Scheduling Timezone**.
 - e. Select **Create & Run Now**.

After your transfer has run, you can see the results of your transfer in **Data Workbench > Databases**.

Further Information

- [List of Options for PostgreSQL Data Connector](#)