

## 8.5.4

## Load the Data

To complete the ETL process, one last phase remains—the load phase. Specifically, we need to load the data into a database. To do so, we'll use pgAdmin to import our saved CSV files.

In the pgAdmin window, select the drop-down menu for the `crowdfunding_db` database. Before we import the CSV files, we need to determine the order in which to do so.



### REWIND

If a table has a Foreign Key constraint, we must first import the tables that this constraint references.

The `campaign` table has three foreign keys: "contact\_id", "category\_id", and "subcategory\_id". So, we need to import the contact, category, and subcategory CSV files before importing the campaign CSV file.

To import the CSV files, complete the following steps:

1. Right-click the first table (contacts).
2. On the menu that appears, scroll to Import/Export.
3. Toggle the button to show Import.
4. Click the ellipsis in the Filename field to search for your project folder.
5. Select contacts.csv. Make sure that Format is set to "csv" and that Encoding is blank.
6. Leave the OID field as is, but toggle the Header field to "Yes," and select the comma as the delimiter.
7. Click OK to begin importing the data.

If the import succeeds, a dialog box, named "Import - Copying table data," will appear at the bottom of the pgAdmin page and contain a "Successfully completed" message.

## Import - Copying table data



Copying table data 'public.contacts' on database 'crowdfunding\_db' and server (localhost:5433)

Tue Jan 11 2022 16:32:28 GMT-0600 (Central Standard Time)

0.04 seconds

More details...

Stop Process



Successfully completed.

Check the import by typing `SELECT * FROM contacts;` in the query editor, and run the query. The resulting table should mirror the CSV file, as the following image shows:

	contact_id [PK] integer	first_name character varying (50)	last_name character varying (50)	email character varying (100)
1	4661	Cecilia	Velasco	cecilia.velasco@rodrigues.fr
2	3765	Mariana	Ellis	mariana.ellis@rossi.org
3	4187	Sofie	Woods	sofie.woods@riviere.com
4	4941	Jeanette	Iannotti	jeanette.iannotti@yahoo.com
5	2199	Samuel	Sorgatz	samuel.sorgatz@gmail.com
6	5650	Socorro	Luna	socorro.luna@hotmail.com
7	5889	Carolina	Murray	carolina.murray@knight.com
8	4842	Kayla	Moon	kayla.moon@yahoo.de
9	3280	Ariadna	Geisel	ariadna.geisel@rangel.com
10	5468	Danielle	Ladeck	danielle.ladeck@scalfaro.net

Practice importing the data from the remaining CSV files yourself in the following Skill Drill:

## SKILL DRILL

Import the data from each remaining CSV file into its corresponding table. Then check the import by using a SQL query. If you need to re-create a table, remember to first use the **DROP TABLE** statement.

Congratulations on finishing the load phase of the ETL process!

#### NOTE

ETL is the most common way to create a data pipeline—but it's not the only way. The **Extract, Load, and Transform (ELT)** paradigm also exist.

ELT is primarily used for scalable cloud structured and unstructured data. In the ELT process, the data is loaded into the database, which gets transformed when analysis is performed. With ETL, the relational and structured data is transformed before loading the data into a database, which makes analyses easier to perform.

© 2020 - 2022 Trilogy Education Services, a 2U, Inc. brand. All Rights Reserved.