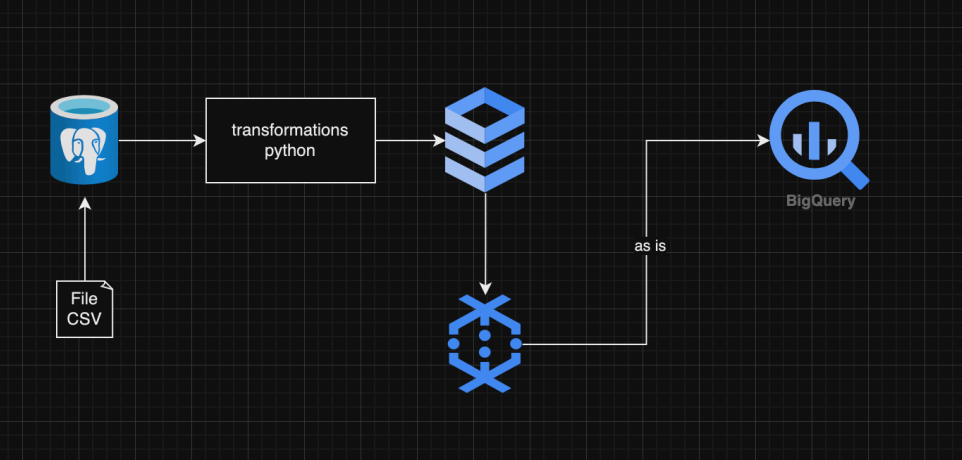
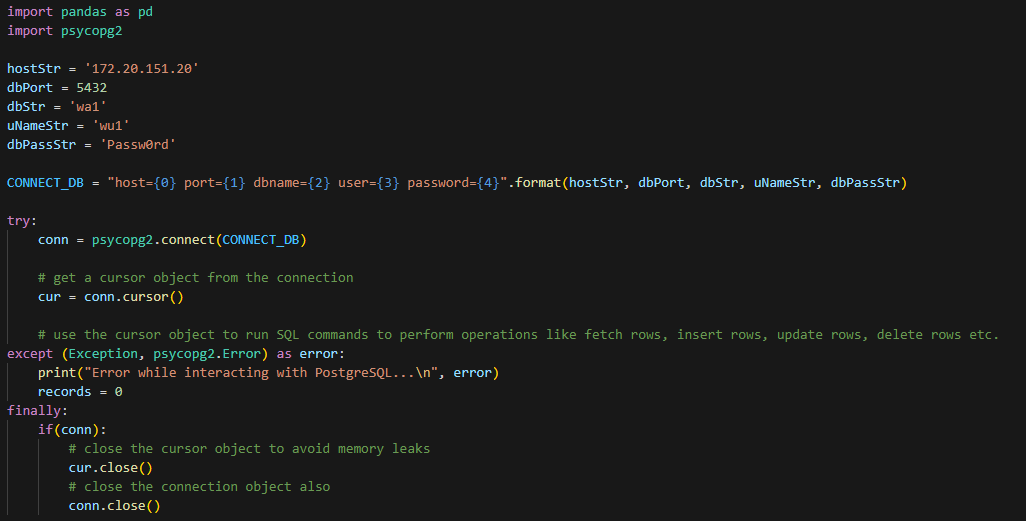
**Flow Data**

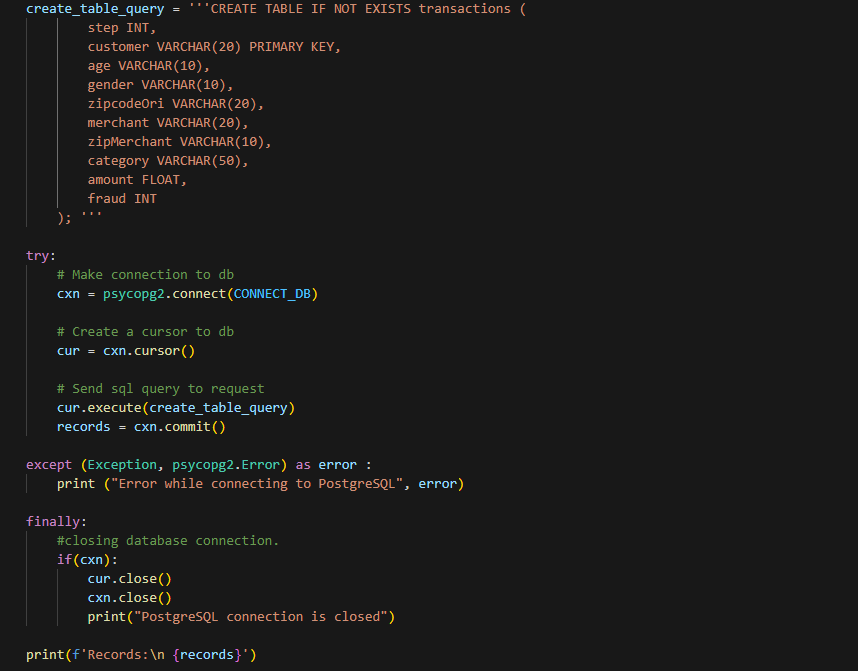


1. Load File CSV ke Postgresql

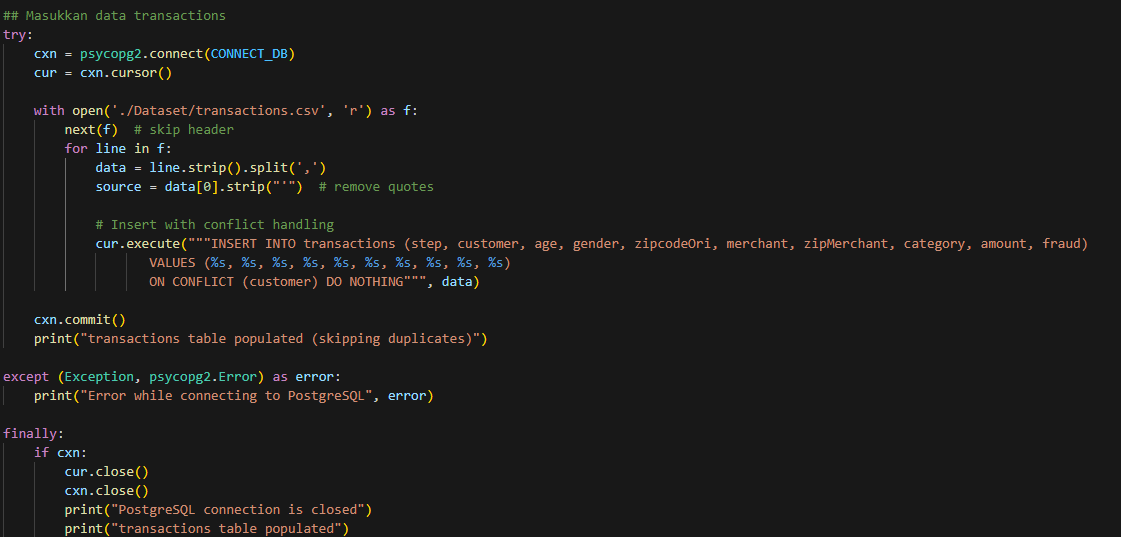
Import library dan declare connection yang akan digunakan.



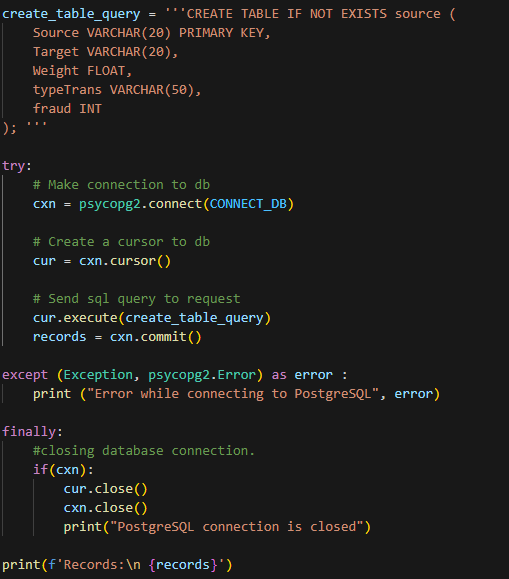
Buat table transactions.



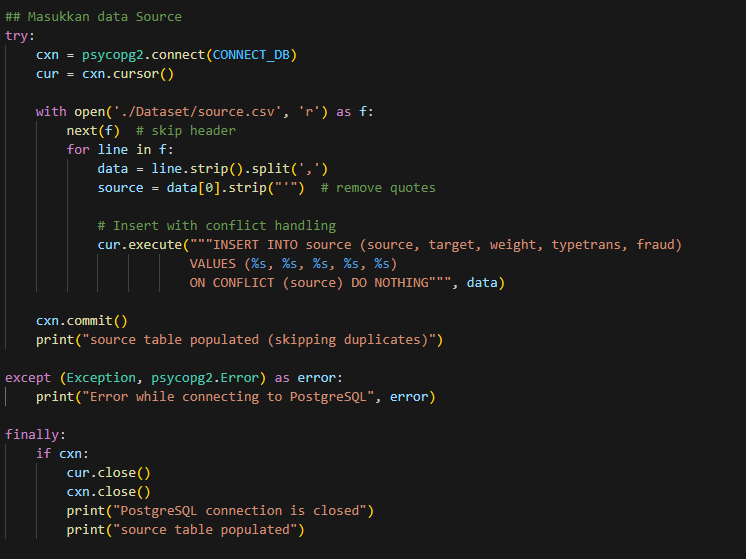
Masukkan data transactions.csv / bs140513\_032310.csv



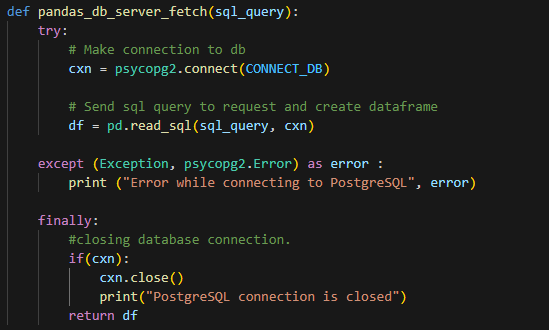
Buat table source

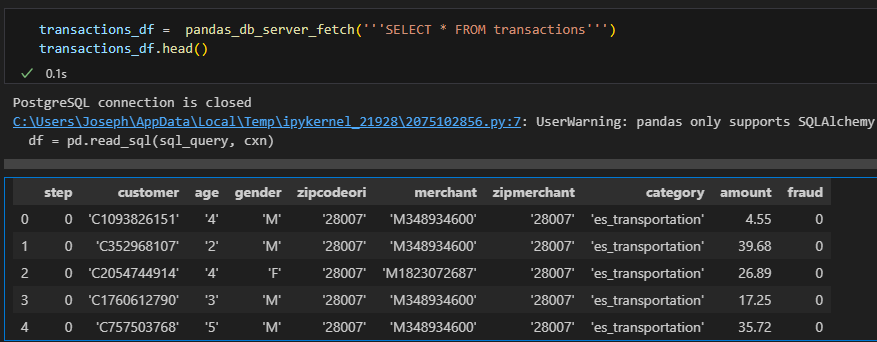


Masukkan data source.csv / bsNET140513\_032310.csv

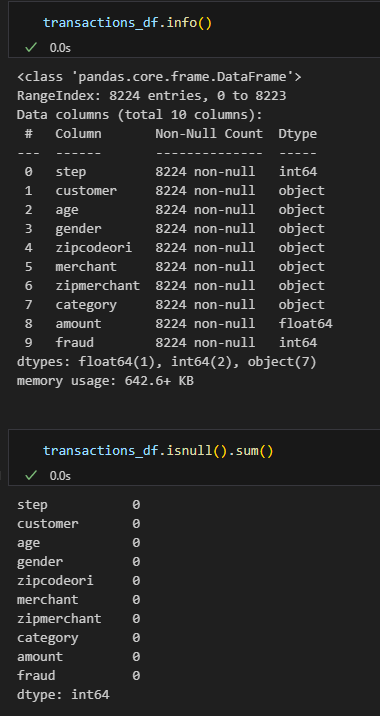


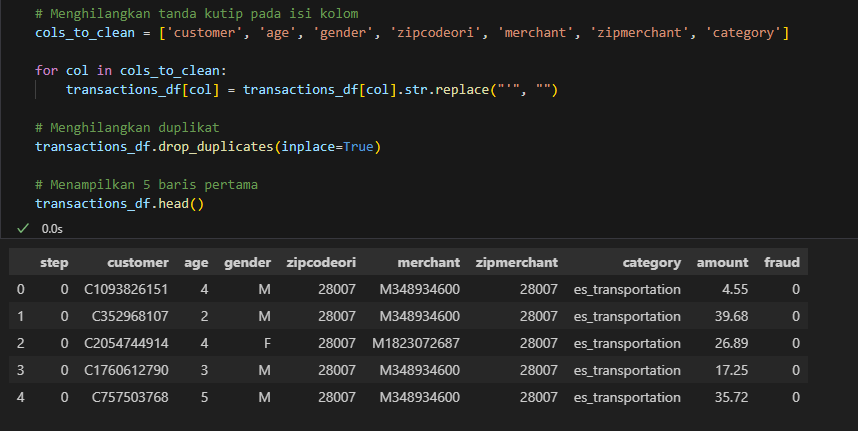
Buat function untuk fetch data.

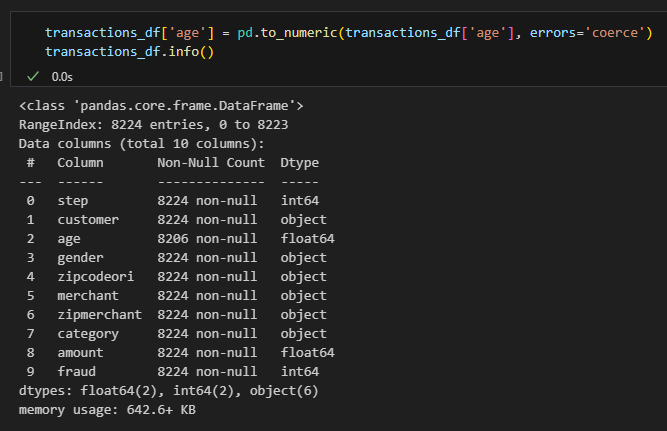


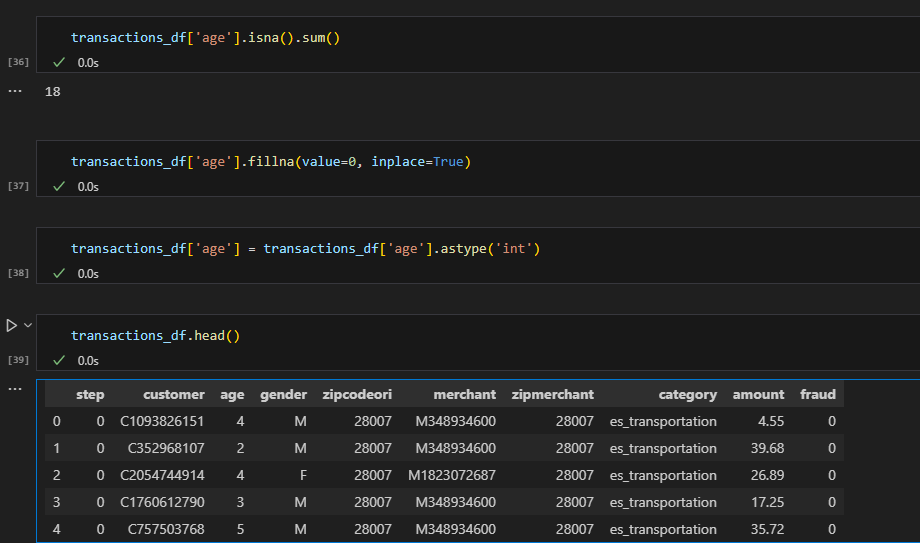


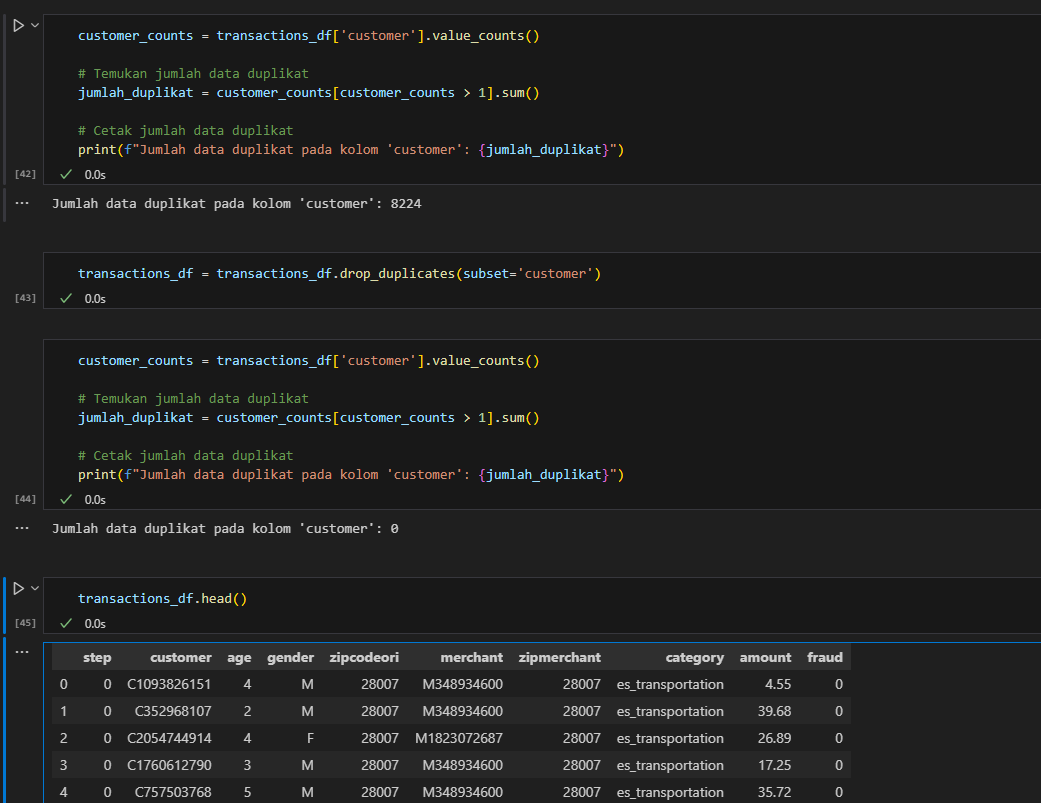
Melihat tipe data dari kolom dan melakukan pengecekan data null.









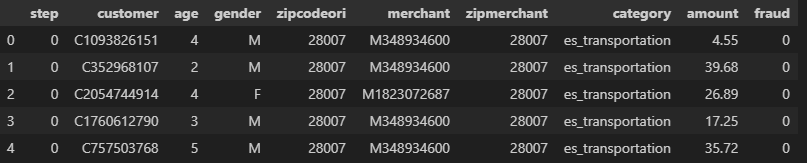


**Cleaning data transaction:**

* Menghilangkan tanda ‘ ‘ pad isi data yang dapat membuat error
* Menghilangkan data-data yang duplikat untuk menghindari data yang tidak akurat.
* Mengubah data age dari string ke int untuk keperluan analisis.

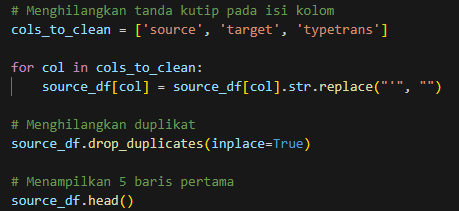
Data Awal

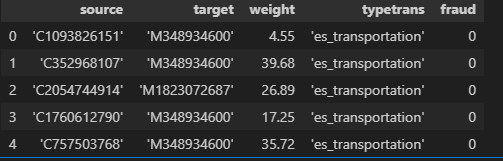


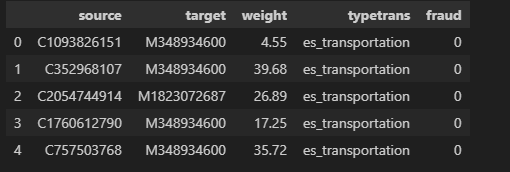
Setelah cleaning

**Cleaning data source:**

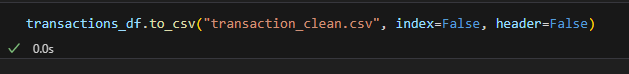
* Menghilangkan tanda ‘ ‘ pad isi data yang dapat membuat error
* Menghilangkan data-data yang duplikat untuk menghindari data yang tidak akurat.

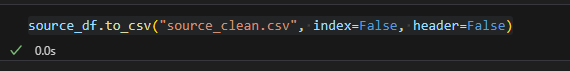






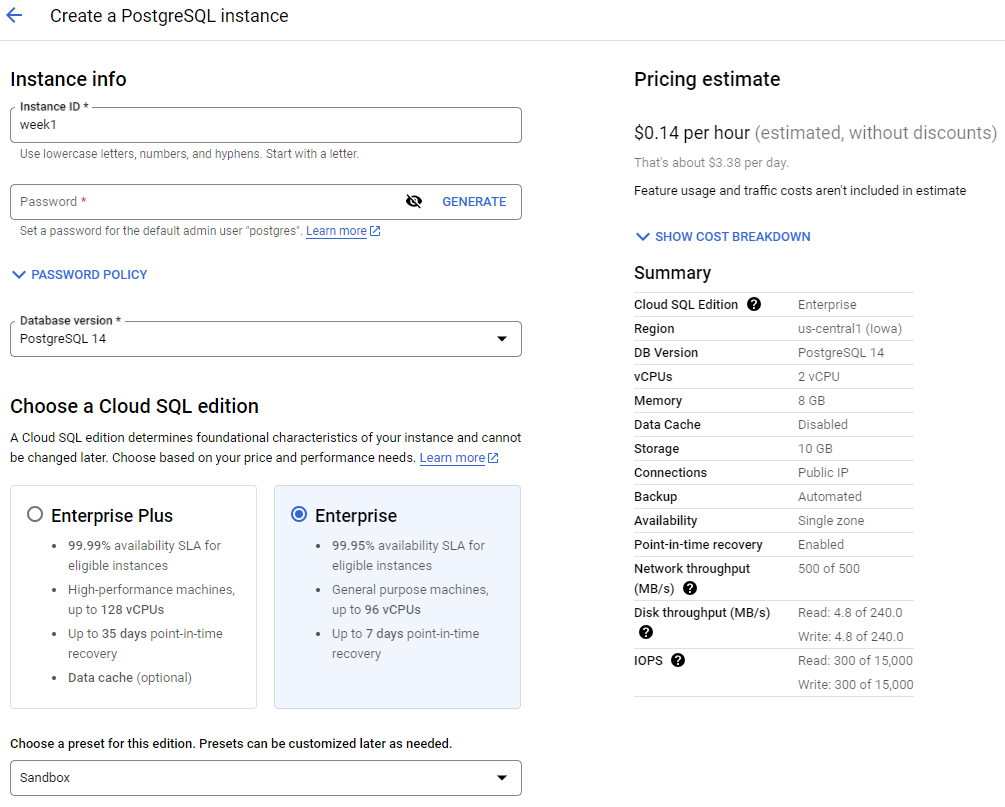
Mengekspor data yang telah dibersihkan kedalam bentuk csv



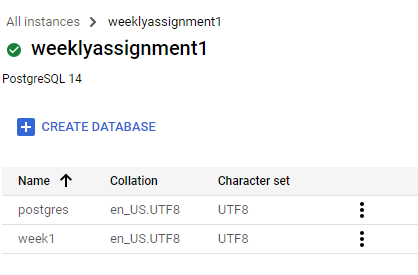


1. Memasukkan data ke CloudSQL

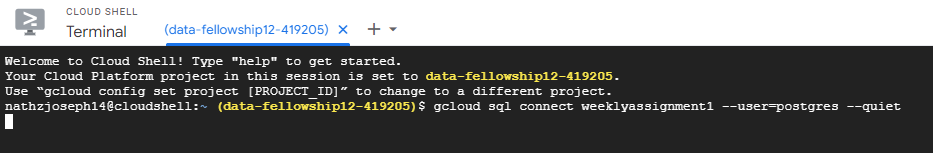
Buat Postgresql Instance baru

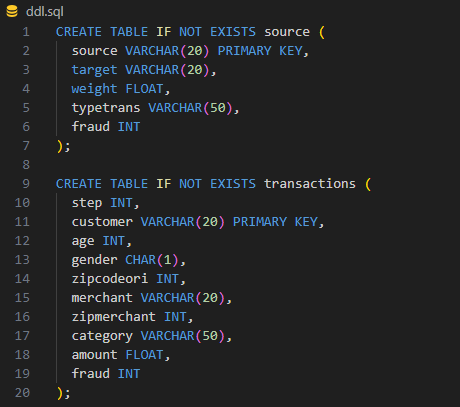


Buat database pada instance

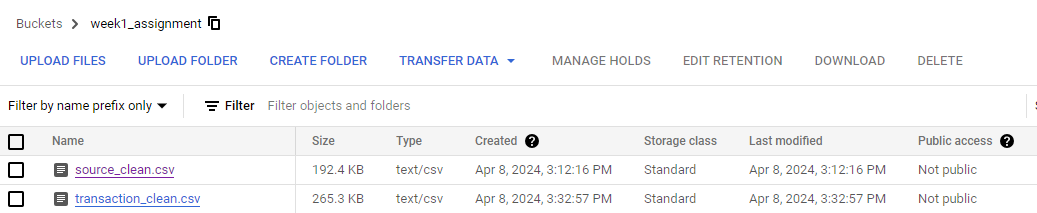


Buka cloud shell dan buat table sebagai tempat untuk melakukan import data yang telah dicleaning.

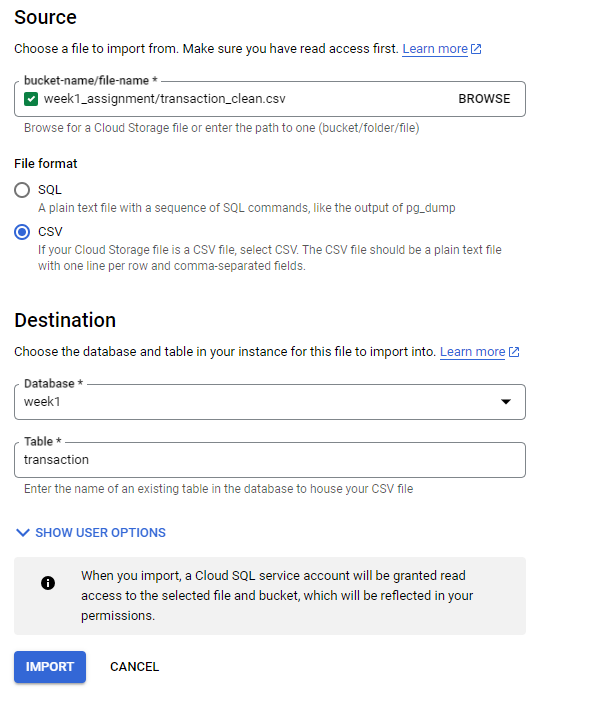




Masukkan data yang akan diimpor kedalam bucket



Import data dari bucket ke database



Pastikan data sudah masuk kedalam database.

