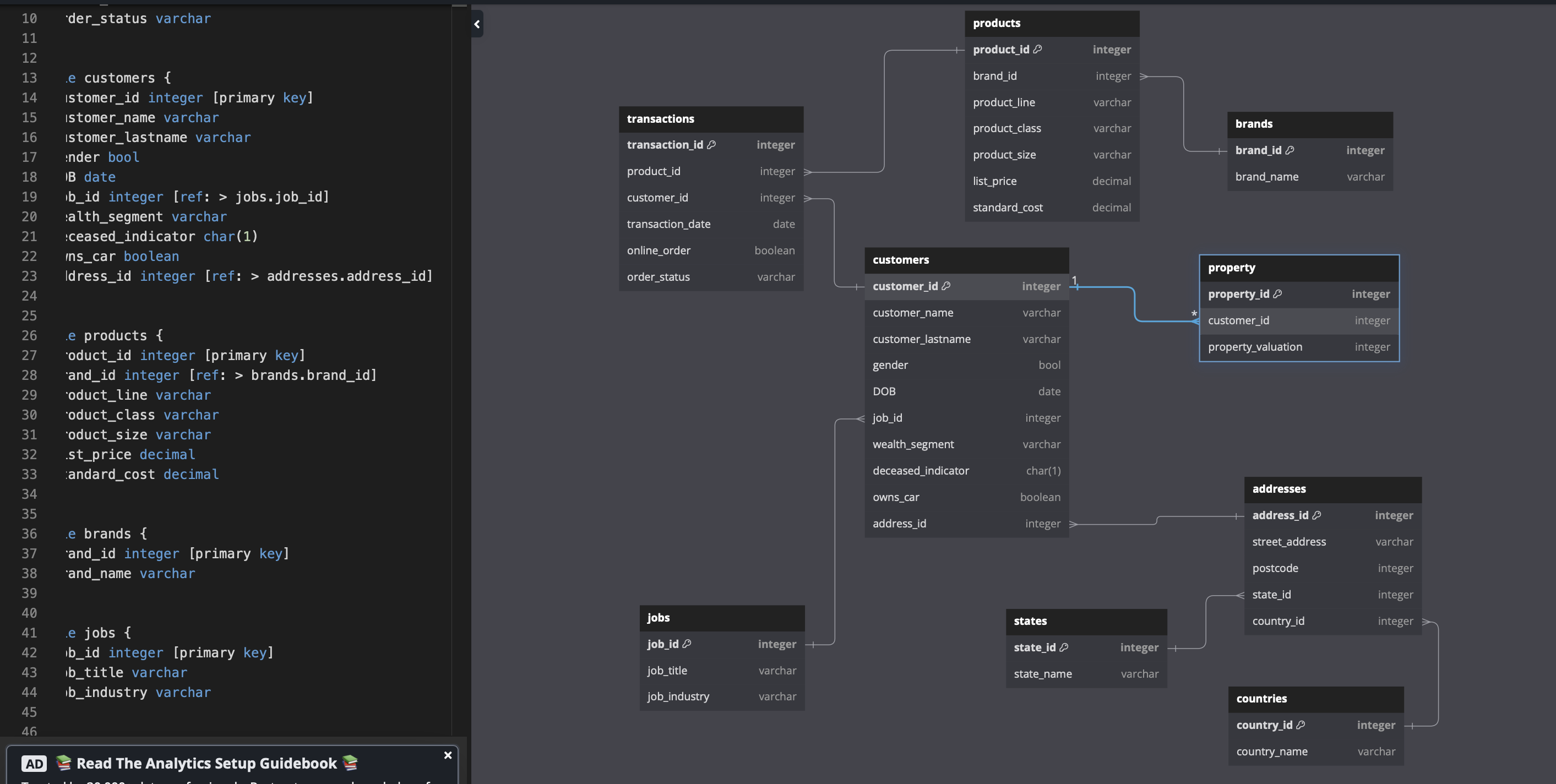
Анализ, почему оригинальная таблица не в нормальной форме:

1. **1НФ Первая нормальная форма** не нарушена. Значения в ячейках атомарные, в одной ячейке только одно значение
2. **2НФ Вторая нормальная форма** нарушена. Если частичные зависимости между колонок. Например, в колонке brand хранится название бренда для каждого товара, но товары с одним и тем же product\_id повторяются для разных транзакций. Это означает, что название бренда дублируется в разных строках, что нарушает 2НФ. В этом случае, необходимо создать отдельную таблицу products, где каждый product\_id будет связан с brand\_id, product\_line, product\_class и другими характеристиками. В transactions оставить только product\_id, без повторяющейся информации о товаре.

3. **3НФ Третья нормальная форма** нарушена, есть транзитивные зависимости. Например, В таблице transactions колонка brand зависит от product\_id, а не от самой транзакции. Это транзитивная зависимость, потому что бренд не является характеристикой самой транзакции.

Для приведения таблицы в нормальную форму была выбрана следующая структура:



В исходной таблице все адреса клиентов указаны только в одной стране, Австралии. Таблицы state и countries вынесены в отдельные на случай, если будут добавляться данные из других стран.

Далее была создана структура таблиц в DBeaver через SQL скрипт.

Пример создания таблицы:

**CREATE** **TABLE** products (

product\_id **INTEGER** **PRIMARY** **KEY**,

brand\_id SERIAL **not** **null** **REFERENCES** brands(brand\_id), -- Auto-increment

product\_line **VARCHAR**,

product\_class **VARCHAR**,

product\_size **VARCHAR**,

list\_price **DECIMAL**,

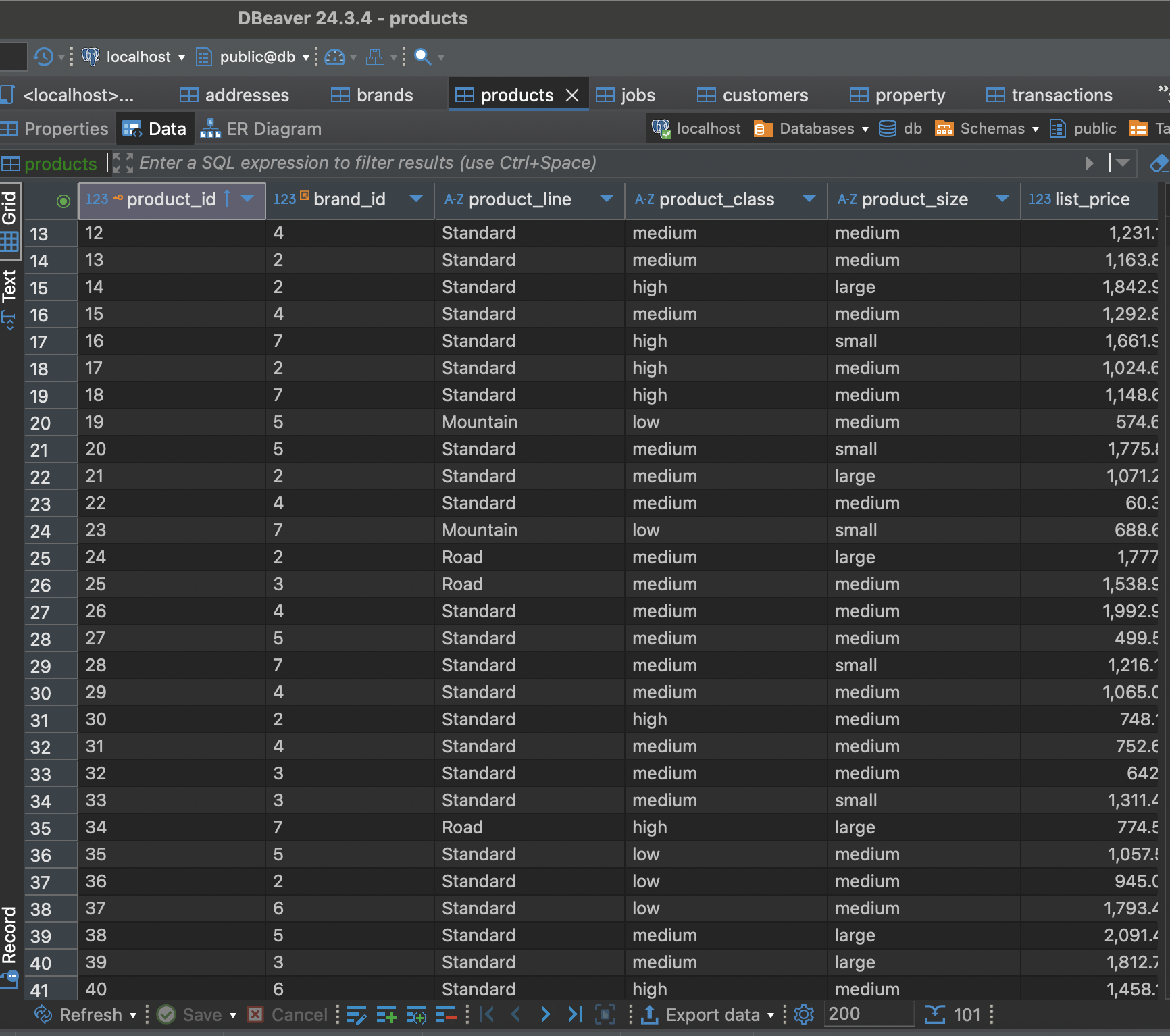
standard\_cost **DECIMAL**

);

Остальные таблицы в документе Script-create table.

Данные были загружены во все таблицы через python код и библиотеку psycopg2.

Код в ноутбуке **homework1 load data.ipynb**

****