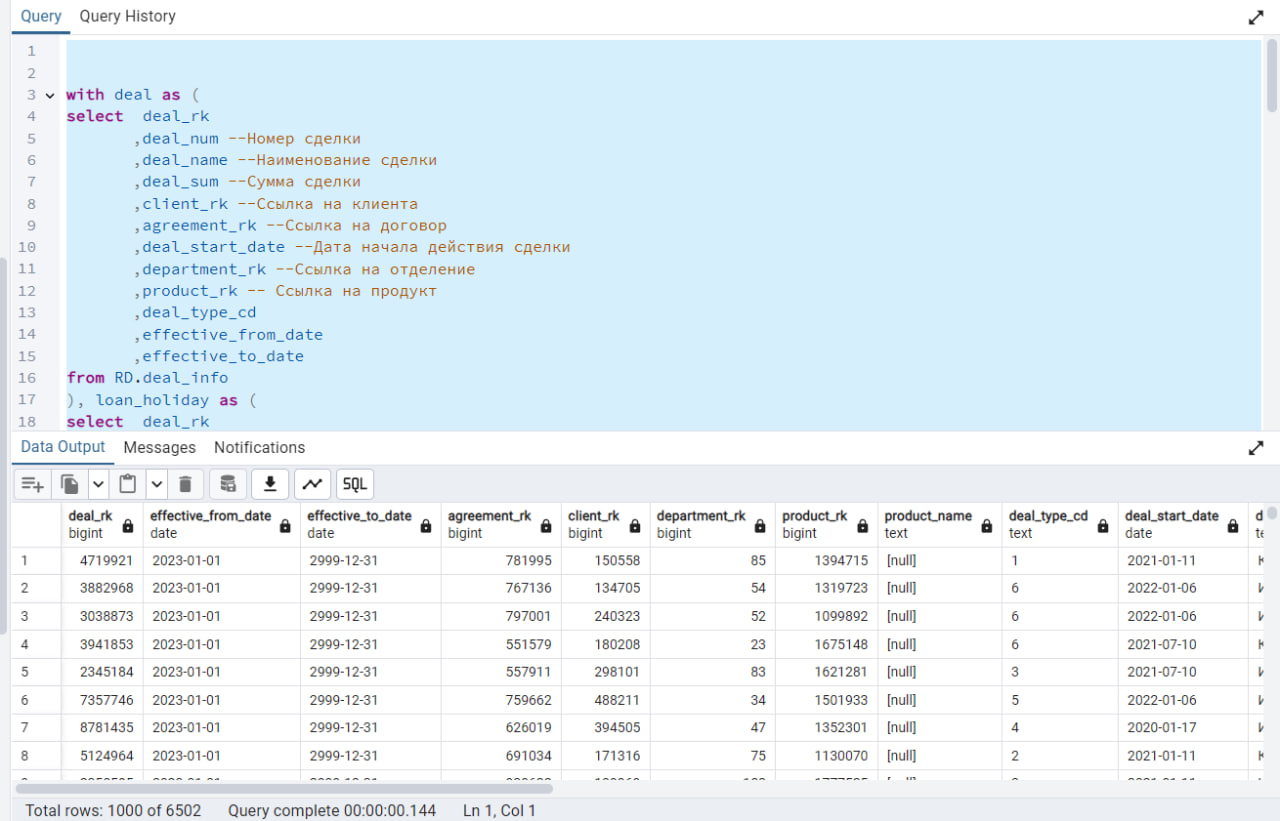
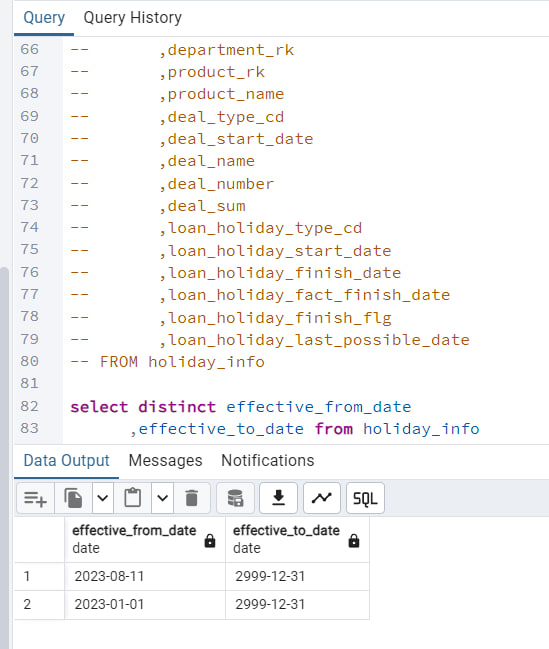
**Ход анализа**

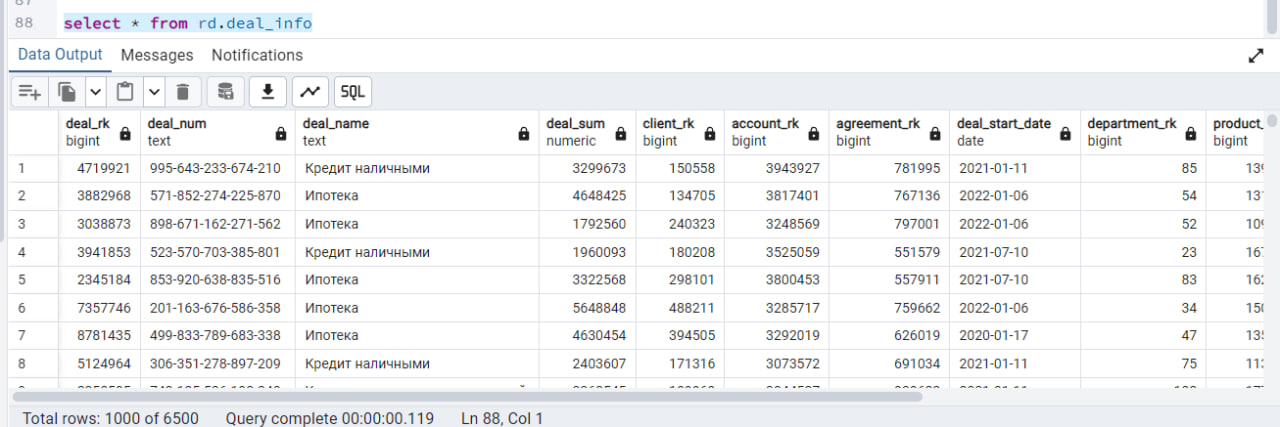
Изначально в прототипе витрины dm.loan\_holiday\_info содержалось 6502 строки:  


И присутсвовали даты 2023-08-11, 2023-01-01 в effective\_from\_date:

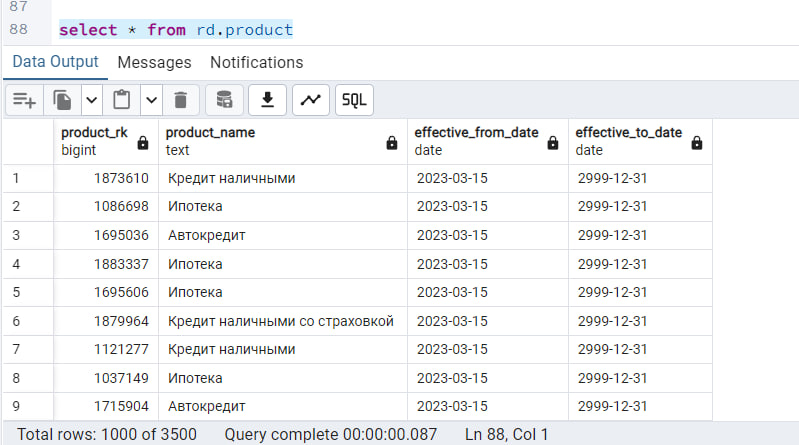


*Слой rd*

количество строк изначально на слое rd deal\_info – 6500

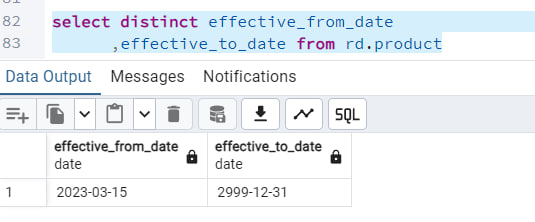


изначально количество строк на слое rd product – 3500

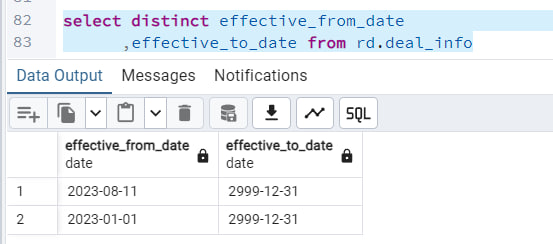


Даты:

В rd.product содержатся данные за дату 2023-03-15



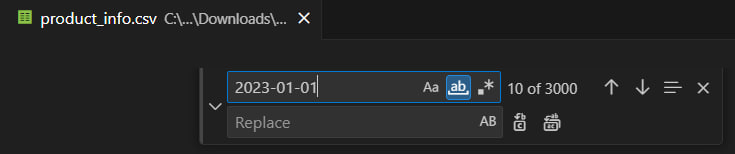
В rd.deal\_info содержатся данные за даты 2023-08-11, 2023-01-01



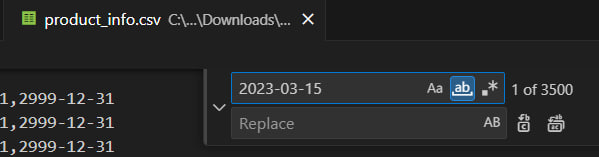
*CSV-файлы*

в файле csv product\_info 10000 строк из которых:

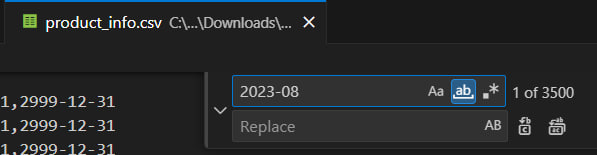
3000 за 01-01-2023



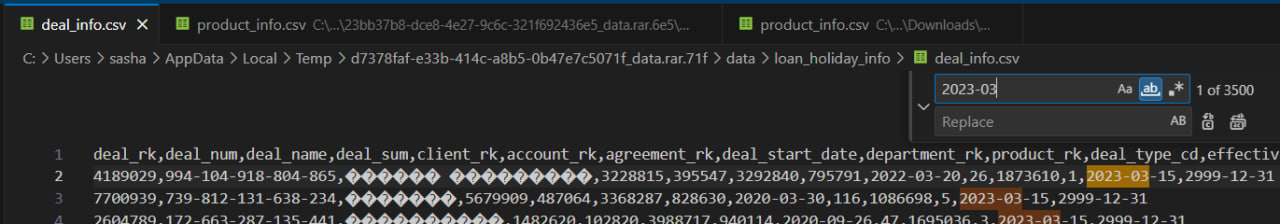
3500 за 03-15



3500 за 08-11



в csv файле deal info всего 3500 строк и все они за 03-15



**Вывод:**

таким образом мы понимаем что на слое rd:

1) в таблице deal\_info отсутсвуют данные за 03-15(которые находятся в csv), при этом есть данные за 08-11 и 01-01

2) в таблице product отсутсвуют данные за 08-11, 01-01, при этом есть данные за 03-15, а в csv файле есть данные за все три даты

для таблицы deal-info мы выбираем полную выгрузку данных из csv, для таблицы product мы выбираем тоже полную выгрзку данных из csv

для первой таблицы очень важно не очищать таблицу перед загрузкой

вторую таблицу будем очищать перед зпгрузкой для избежания дубликатов строк за дату 2023-03-15