**Планы запросов в GreenPlum и PostgreSQL и их оптимизация**

Подключение к инстансу GreenPlum (см. рисунок 1-3).

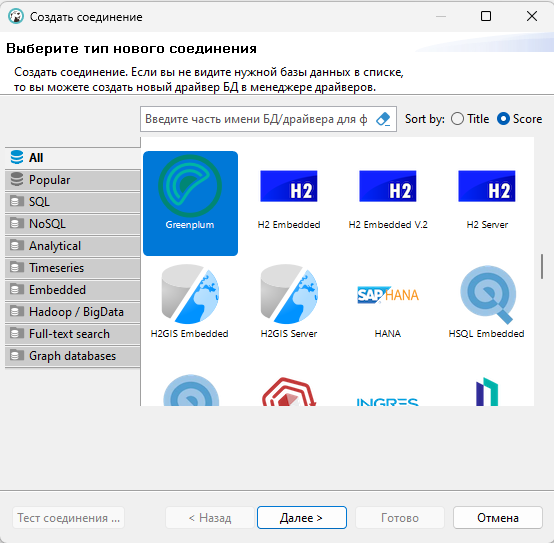


Рисунок 1.

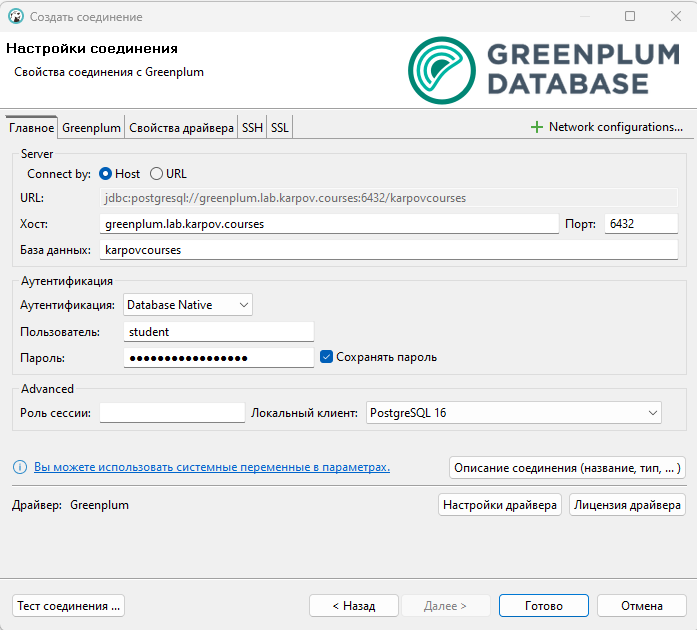


Рисунок 2.

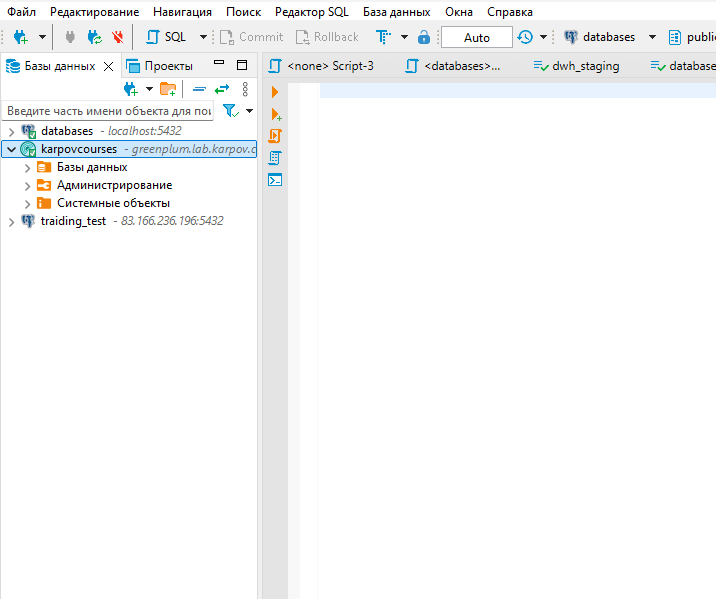


Рисунок 3.

Выполним объединение по ключу покупателя тыблиц заказов и покупателей (см. рисунок 4).

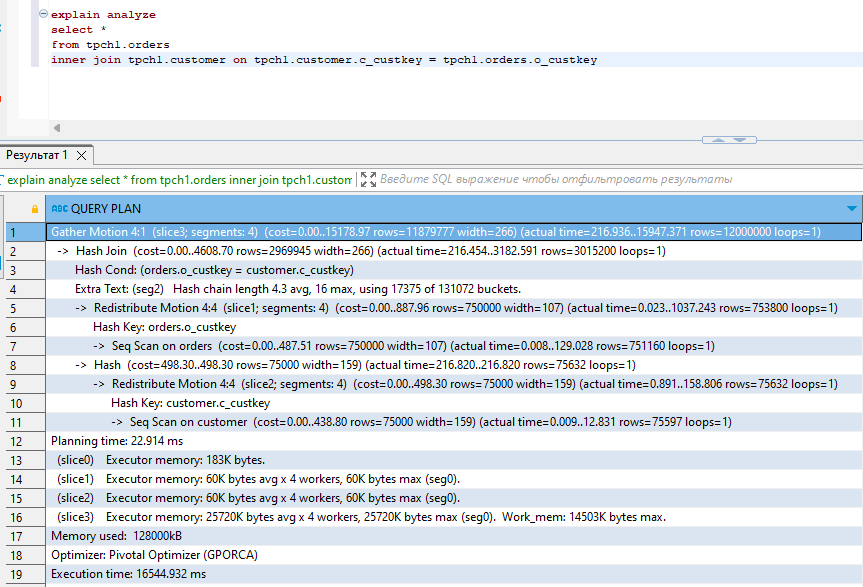


Рисунок 4.

По плану видим, что ожидаемое количество строк ~12.000.000.

Далее считаем таблицу покупателей и посмотрим сколько строк планирует считать оптимизатор с каждого сегмента БД, в которых находятся данные этой таблицы (см. рисунок 5).

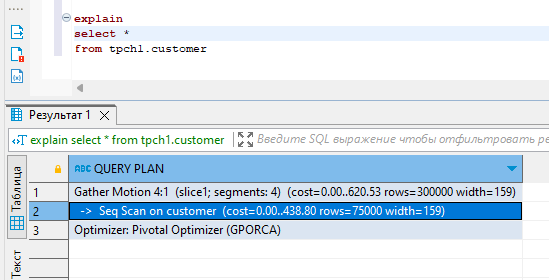


Рисунок 5.

Выполним запрос для другой таблицы с клиентами и сравним планы (см. рисунок 6).

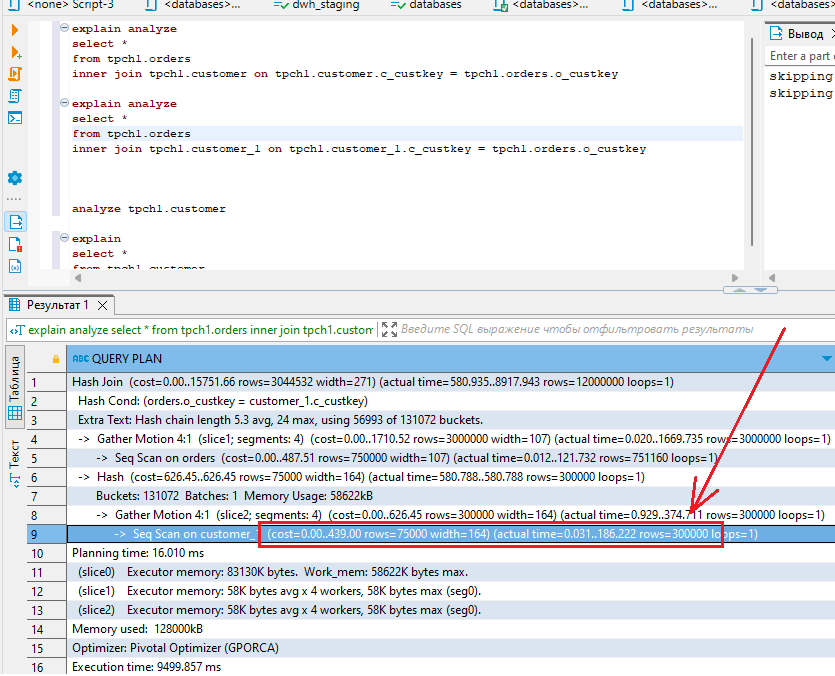


Рисунок 6.

Мы видим, что один из сегментов таблицы с клиентами customer\_1 перекошен в большую сторону в фактическом результате отработки плана запроса => данные распределены неравномерно между сегментами.

Также с помощью функции gp\_dist\_random мы можем посмотреть на распределение данных между сегментами (см. рисунок 7).

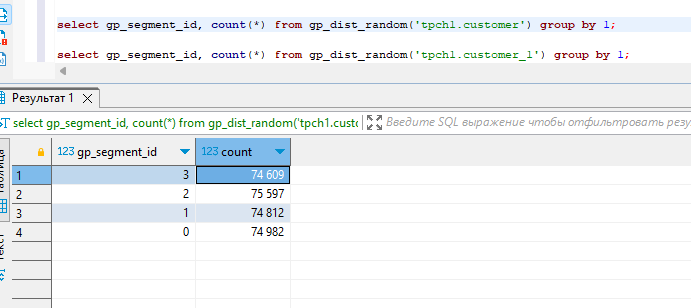


Рисунок 7.

Посмотрим на статистику по внешним таблицам (см. рисунок 8).

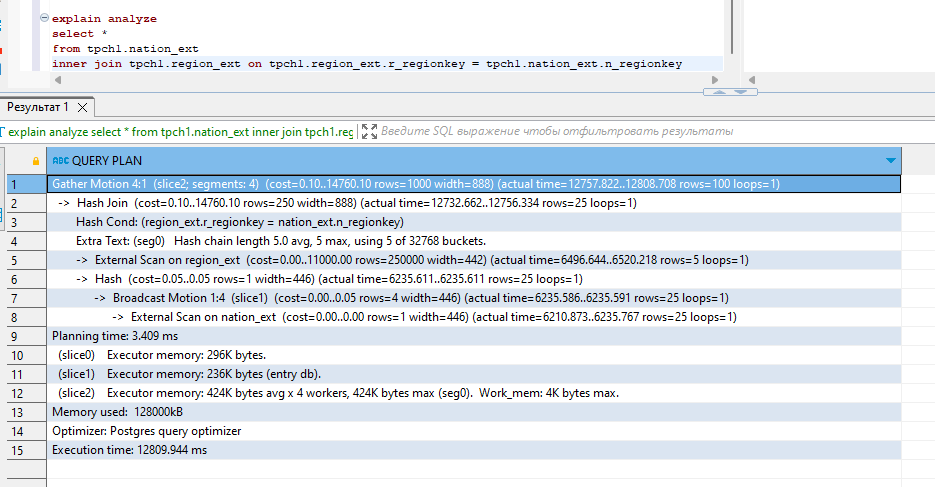


Рисунок 8.

Как видим, статистики нет в плане запроса по этим таблицами, а значения подставляются случайным образом, основываясь на данных о ширине.