Nama: Karno

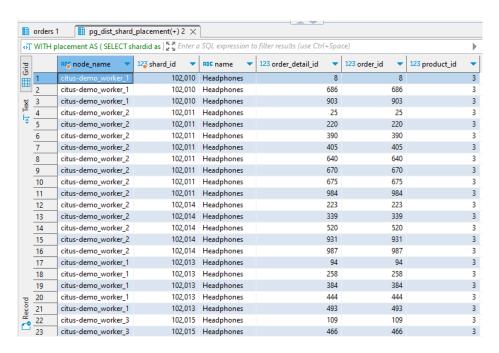
Jurusan: Data Engineer 01

- 1. Oerbedaan replication dan sharding:
 - Replication adalah strategi dimana data disalin atau didistribusikan ke beberapa lokasi/node atau worker untuk tujuan backup data. Ini berguna apabila salah satu worker mati maka ada worker lain yang melakukan backup. Pada replication masing-masing node/worker memiliki data yang sama dengan node/worker lainnya.
 - Sharding adalah strategi dimana data dibagi menjadi beberapa bagian dan dimasukkan kedalam node/worker yang berbeda. Ini bertujuan untuk mempercepat pembacaan data. Pada sharding data yang ada di node/worker yang satu berbeda dengan yang ada di node/worker lain.

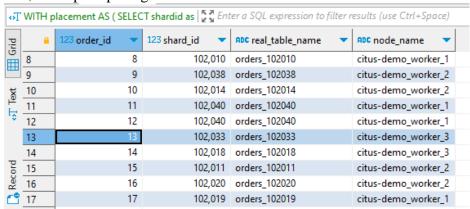
2. Melakukan percobaan untuk membuat referensi table

```
5432 | default | t
                                    2 | citus-demo_worker_1 |
select * from orders_102016 order by order_id limit 5
order id | user id | total price |
                                                created at
                             509.27 | 2023-11-04 12:39:42.770383+00
                             158.35 | 2023-11-04 12:39:42.770383+00
       53
                   2 |
                             514.24 | 2023-11-04 12:39:42.770383+00
985.70 | 2023-11-04 12:39:42.770383+00
      112
      163
                              72.44 | 2023-11-04 12:39:42.770383+00
      208
(5 rows)
WORKER citus-demo worker 2
SELECT count(1) FROM orders
count
 2000
(1 row)
select * from orders_102011 order by order_id limit 5
order_id | user_id | total_price |
                                                 created at
                              43.53 | 2023-11-04 12:39:42.770383+00
       15 |
                   1 |
                             134.41 | 2023-11-04 12:39:42.770383+00
```

3. Di node/worker mana sajakah product headphone tersimpan, tunjukkan shard_id nya Product headphone tersimpan di semua node dengan shard_id seperti digambar



4. Di node/worker mana saja order dengan id 13 tersimpan Order dengan id 13 tersimpan di node "citus-demo-worker_3" dengan shard_id 102,033 seperti pada gambar



- 5. Replication sebaiknya digunakan Ketika kita menjalankan suatu system yang harus terus menyediakan data tanpa henti. Sehingga Ketika salah satu server atau system mengalami error maka server yang lain siap digunakan tanpa harus mengganggu data utama. Replication bertindak dengan cara menyediakan data cadangan supaya siap digunakan Ketika dibutuhkan.
- 6. Sharding sebaiknya digunakan Ketika kita memiliki dataset yang sangat besar atau Ketika computer memiliki beban kerja yang tinggi. Sharding bekerja dengan cara membagi data ke beberapa server (shard) sehingga masing-masing shard dapat mengatasi Sebagian masalahnya.