

Таблиця review:

Ідея цієї таблиці, те що це основна таблиця, тобто та яка містить всі дані і опирається на review_id так як воно є єдиним унікальним у нашій базі даних (customer_id та product_id можуть повторюватись). В primary keys цієї таблиці я додала product_id, star_rating, review_id - перші два для того аби я змогла виконати перші два запити за допомогою цієї таблиці, третій чисто ідейно щоб ми могли глянути на конкретний рядок, ролі він тут не грає.

Таблиця customer_review:

Ця таблиця має ті ж дані що й попередня, різниця лише в primary keys. Так як в cassandra досить обмежена кількість команд для queries для 3 та 5 запиту не було виходу окрім створити окрему таблицю. Тут primary keys так само взяті для виконання запитів.

Таблиця customer_statistics:

Моїм основним підходом було створення мінімальної кількості таблиць, а таблиць відповідно з мінімальною кількістю даних, адже для великих баз даних в цьому полягає оптимальність. Для 4 query не можливо було обійтись без додаткової таблиці (так як джойн і тд касандра не підтримує), через це я створила таблицю з мінімальною кількістю стовпців, кількість рядків в цій таблиці теж буде менша ніж в попередніх і буде відповідати кількості клієнтів які залишали хоча б один відгук. Я назвала цю таблицю статистика покупців щоб надати їй трохи більш логічного сенсу, і в ній зберігається список всіх продуктів на які залишив відгук користувач і загальна кількість відгуків.

Результат запуску запитів на тестових даних можна подивитись в screenshots фолдері.