

Hive

Представьте, что подготовлены требования на заполнение таблицы следующей структуры: id bigint, col1 string, col2 string, col3 string. При этом таблица должна быть партиционирована по полю p int. Формат файлов - parquet.

- Создайте таблицу в hive с указанной выше структурой.
- Создайте тестовые данные в таблице: 10 партиций от 1 до 10, с 1000 строчками в каждой. Поле id заполняется любым уникальным числом, поля col1, col2, col3 - 3 случайных алфавитно-цифровых значения (f9x, xcf, 123 и т.д.). Подсказка: для генерации тестовых данных можно использовать конструкцию lateral view posexplode, встроенные функции rand(), chr().
- Сделайте несколько выборок: для каждой партиции выведите минимальное, максимальное значение id и общее количество строк для случаев когда
 - все три текстовых поля заполнены буквами
 - все три текстовых поля заполнены цифрами

Представьте, что через некоторое время служба безопасности проверила нашу таблицу и потребовала удалить из таблицы поле col2, так как оно не удовлетворяет правилам безопасности.

- Удалите из таблицы поле col2, при этом данные по остальным полям в таблице должны быть сохранены. Подсказка: напрямую поле из таблицы удалить не получится, нужен ряд последовательных действий.