

| | |
|-------|-----------------|
| Nama | Utari W Ardhana |
| Kelas | DE 1 |

Analisa Big Data Menggunakan Hive

Download file csv yang dilampirkan untuk membaca informasi pada data yang akan dianalisa di hive, data yang didownload memiliki metadata sebagai berikut:

- **TransactionNo** (kategorikal): nomor unik enam digit yang mendefinisikan setiap transaksi.
- **Date** (numerik): tanggal setiap transaksi dibuat.
- **ProductNo** (kategorikal): karakter unik lima atau enam digit yang digunakan untuk mengidentifikasi produk tertentu.
- **Product** (kategorikal): nama produk/item.
- **Price** (numerik): harga setiap produk per unit dalam pound sterling (£).
- **Quantity** (numerik): kuantitas setiap produk per transaksi. Nilai negatif terkait transaksi yang dibatalkan.
- **CustomerNo** (kategorikal): nomor unik lima digit yang mendefinisikan setiap pelanggan.
- **Country** (kategorikal): nama negara tempat tinggal pelanggan.

1. Buatlah perintah SQL untuk membuat tabel berdasarkan dataset yang diberikan (transactions.csv)

Jawaban :

```
CREATE TABLE transaction (  
    TransactionNo INT,  
    Date DATE,  
    ProductNo INT,  
    Product VARCHAR(30),  
    Price INT,  
    Quantity INT,  
    CustomerNo INT,  
    Country VARCHAR(30)  
)  
ROW FORMAT DELIMITED  
FIELDS TERMINATED BY “,”  
STORED AS TEXTFILE;
```

2. Buatlah perintah Hive untuk melakukan load data pada tabel yang diberikan dengan asumsi file diupload pada direktori “/user/msib/”

Jawaban :

```
LOAD DATA INPATH "USER/uut/transaction.csv" OVERWRITE INTO TABLE  
uut_db_transaction
```

3. Buat view tabel untuk membuat reporting berupa

a. Top 10 revenue berdasarkan ProductName

Jawaban :

b. Top 10 quantity berdasarkan Quantity

Jawaban :

```
SELECT  
    Description  
    SUM(quantity) AS top_quantity  
FROM uut_db_transaction  
GROUP BY Description  
ORDER BY SUM(quantity) DESC  
LIMIT 10;  
  
#View table  
SELECT * FROM uut_db.summary_top_quantity;
```

Q: Bagaimana cara mengerjakan tugas ini?

A: Download file word ini, isi jawaban hanya berupa perintah SQL yang dapat dijalankan di Hiv