

Data Analytics Pipeline:

SQL Assignment

Instructions:

1. Make a new file to answer the question.
2. Named your file with: Fullname_SQL.
3. Save as PDF.
4. You may discuss with your friend but do not copy paste your work.
5. Submit your answer in google classroom before January 17th 2022, 23:59 WIB.

Questions:

Part I

Dataset: Data Pelaporan Kedatangan Penduduk (Data dapat diunduh di google classroom dengan nama 'Data Pelaporan Kedatangan Penduduk.csv').

Jawablah pertanyaan berikut disertai dengan query yang sesuai dan tampilkan sample output tabelnya.

1. Berdasarkan kota/kabupaten, urutkanlah jumlah kedatangan pendatang baru dari yang paling banyak
2. Tentukanlah 10 kecamatan dengan jumlah kedatangan pendatang baru terbanyak dan tersedikit
3. Manakah pendatang baru yang lebih banyak: laki-laki atau perempuan?
4. Untuk setiap kota/kabupaten administrasi, berapa jumlah kecamatan dan kelurahan dan jumlah pendatang baru?

Part II

Dataset: Indonesia Patient Covid Case (Data dapat diunduh di google classroom dengan nama indo_covid.zip.)

1. **Buatlah script SQL INNER JOIN** untuk menampilkan jumlah acc_confirmed dan acc_negative dari table case_csv dan positive_total dan odp_total dari table jabar_csv dengan menggunakan date sebagai field untuk melakukan JOIN terhadap dua table (case_csv dan jabar_csv). **Lalu tampilkan juga output querynya.**

2. Jelaskan maksud dari query SQL berikut:

```
SELECT          jabar_csv.`date`,          cases_csv.acc_confirmed,  
cases_csv.acc_negative, jabar_csv.positive_total, jabar_csv.odp_total  
FROM jabar_csv LEFT JOIN cases_csv ON jabar_csv.`date` =  
cases_csv.`date` WHERE cases_csv.acc_confirmed IS NOT NULL ORDER BY  
cases_csv.acc_negative DESC;
```