

Test Data Engineer

Studi Kasus 1: Analisis Waktu Bongkar Muat Kapal

Deskripsi:

Pelabuhan ingin menganalisis efisiensi operasi bongkar muat kapal. Data operasi memiliki atribut: ship_id, arrival_time, departure_time, cargo_weight, crane_id, operation_time.

Tugas Kandidat:

1. Bangun pipeline untuk memproses data operasi bongkar muat yang dikirimkan dalam format CSV harian.
2. Transformasikan data untuk:
 - a. Menghitung waktu rata-rata bongkar muat per kapal.
 - b. Mengidentifikasi crane dengan efisiensi terendah berdasarkan waktu operasi.
3. Simpan hasil analisis ke dalam database PostgreSQL.
4. Buat laporan mingguan dalam format PDF atau dashboard yang menunjukkan:
 - a. Kapal dengan waktu bongkar muat terlama.
 - b. Crane dengan performa terbaik dan terburuk.

Kriteria Penilaian:

- Efisiensi pipeline data.
- Ketepatan hasil analisis.
- Penyajian laporan yang jelas dan informatif.

Studi Kasus 2 : Analisis Kapasitas Kontainer di Pelabuhan Korea Selatan

Persiapan Dataset

Unduh dataset [di sini](#).

Dataset memiliki kolom utama:

- **Port Name:** Nama pelabuhan
- **Year:** Tahun data diambil
- **Month:** Bulan data diambil

- **Capacity:** Jumlah kapasitas kontainer (TEU)

Pertanyaan dan Tugas

1. Baca dataset menggunakan Python (library Pandas).
2. Hitung rata-rata kapasitas kontainer setiap pelabuhan per tahun.
3. Identifikasi pelabuhan dengan kapasitas kontainer tertinggi rata-rata setiap tahun.
4. Buat visualisasi tren kapasitas kontainer bulanan untuk setiap pelabuhan.
5. Simpan hasil analisis ke dalam file CSV bernama `port_analysis_results.csv`.

Studi Kasus 3 : Kapasitas Kontainer dan Informasi Pelabuhan

Deskripsi:

Diberikan dua tabel:

1. **port_container_capacity:**
 - `port_id` (INT): ID pelabuhan.
 - `year` (INT): Tahun data.
 - `month` (INT): Bulan data.
 - `capacity` (INT): Kapasitas kontainer (TEU).
2. **port_details:**
 - `port_id` (INT): ID pelabuhan.
 - `port_name` (VARCHAR): Nama pelabuhan.
 - `location` (VARCHAR): Lokasi pelabuhan.

Tugas:

1. Tulis query untuk menghasilkan nama lokasi, kuartal (Q1, Q2, Q3, Q4), dan total kapasitas kontainer.
2. Tulis query untuk mendapatkan lokasi, rata-rata kapasitas tahunan, dan perbedaan rata-rata kapasitas antara kedua lokasi.
3. Tulis query untuk mendapatkan nama pelabuhan, lokasi, dan bulan-bulan di mana kapasitas meningkat dibandingkan bulan sebelumnya.