

**Muhammad Azhar Rasyad**

Teknik Informatika 1

0110217029

STT Terpadu Nurul Fikri

## **Data Warehouse**

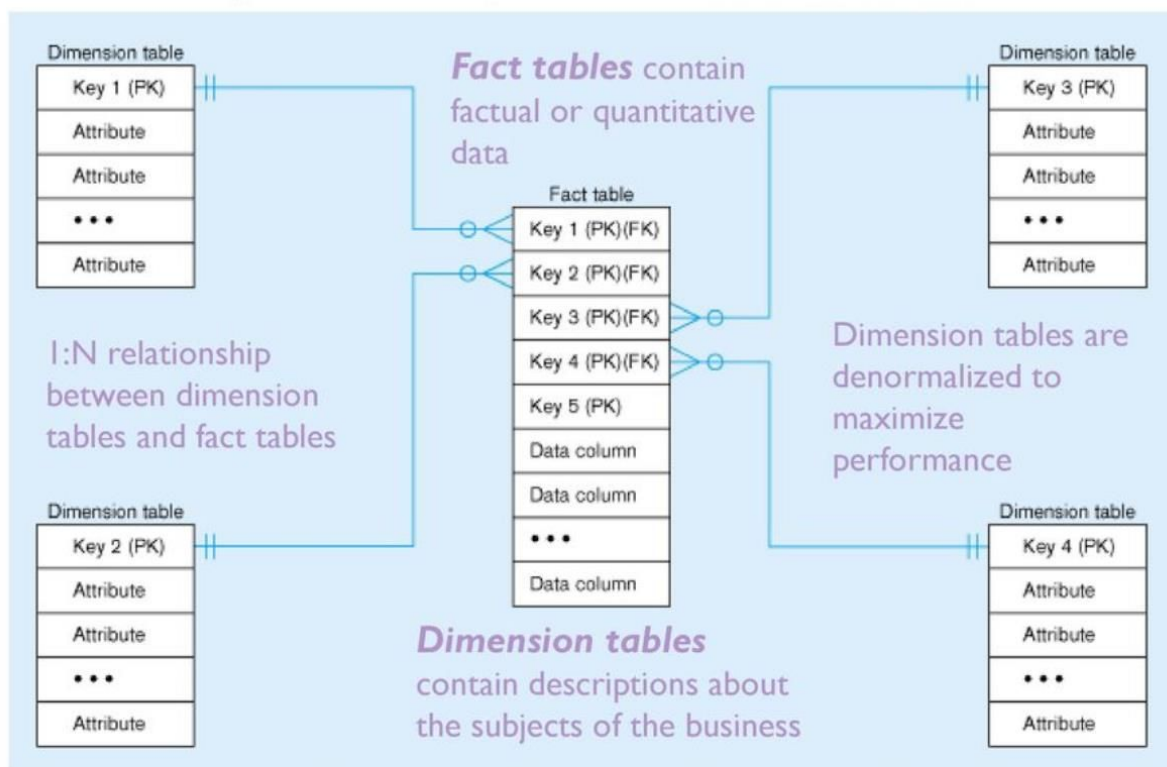
### **Quiz Extraction Transformation Loading**

1. Apa yang dimaksud dengan ETL ?
2. Jelaskan dan berikan lima contoh proses *cleansing* data pada proses transformation !!
3. Apa yang dimaksud skema bintang !

### **Jawaban**

1. **ETL** atau **Extract Transform Load** merupakan langkah di dalam pemrosesan data pada database yang melibatkan kegiatan **extracting** data-data dari sumber data, mempertahankan kualitas data, menerapkan standarisasi data, menyajikan dalam berbagai **transform** untuk kemudian dialirkan atau **load** ke data warehouse.
2. 5 Cleansing Data :
  - a. **Parsing** yaitu Melakukan penguraian untuk menempatkan dan mengidentifikasi elemen data individual dalam file-file sumber dan kemudian melakukan isolasi data ini ke dalam file target.
    - i. Contoh : Parsing data untuk menentukan first name, last name, middle name
  - b. **Correcting** yaitu Melakukan koreksi hasil parsing komponen data individual dengan menggunakan algoritma data yang canggih dan sumber data sekunder.
    - i. Contoh : Perbaikan data field alamat dan kode pos
  - c. **Standardizing** yaitu Penerapan standarisasi format data atau konversi data yang konsisten sesuai aturan bisnis proses organisasi.
    - i. Contoh : Pemberian prefix field nama (gelar depan) atau (gelar belakang), standar data alamat atau jalan

- d. **Matching** yaitu Pencarian dan mencocokkan data dari seluruh data yang diparsing untuk diperbaiki dan memenuhi standar aturan bisnis yang ditentukan untuk menghilangkan duplikasi data.
    - i. Contoh : Identifikasi nama atau alamat yang sama
  - e. **Consolidating** yaitu Melakukan analisa dan identifikasi hubungan antara (relation) data dan melakukan konsolidasi atau penggabungan menjadi satu representasi data.
    - i. Contoh : Konsolidasi data dari dua sumber OLTP menjadi satu representasi data, seperti system CRM & ERP untuk representasi data produk atau jasa atau customer atau employee.
3. Skema Bintang merupakan model data dimensional yang mempunyai *fact table* di bagian tengah, dikelilingi oleh tabel dimensi yang terdiri dari data *reference* (yang bisa di-*denormalized*). Berikut contoh gambar skema bintang :



Excellent for ad-hoc queries,  
but bad for online transaction processing