



## Project Based Internship

# Principles of Data Warehouse & Metadata

Principles of Data Warehouse and Metadata: A Comprehensive Guide

## Daftar Isi

A. Introduction	3
B. Principles of Data Warehouse	3
1. Data Integration	3
2. Historical Data	4
3. Non-Volatile Data	4
4. Query Performance	5
5. Data Access	5
C. Metadata	6
Studi Kasus: Pentingnya Metadata dalam Analisis Perusahaan Ritel	7
REFERENCE	9

## A. Introduction

Data Warehouse adalah proses mengumpulkan, menyimpan, dan mengelola data dari berbagai sumber untuk mendukung proses pengambilan keputusan. Salah satu komponen utama data warehouse adalah metadata, yang memberikan informasi berharga tentang data yang disimpan di data warehouse. Metadata dapat mencakup informasi tentang sumber data, formatnya, strukturnya, dan kualitasnya. Pada artikel ini, kita akan mengeksplorasi prinsip-prinsip data warehouse dan memeriksa pentingnya metadata dalam membangun dan memelihara data warehouse yang baik. Artikel juga akan memberikan contoh bagaimana metadata dapat digunakan untuk meningkatkan proses pengambilan keputusan dan meningkatkan nilai data warehouse.

## B. Principles of Data Warehouse

Data warehouse adalah proses mengumpulkan, menyimpan, dan mengelola data dari berbagai sumber untuk mendukung proses pengambilan keputusan. Ini memungkinkan pengorganisasian data untuk menggabungkan data mereka dari berbagai sumber ke dalam satu tampilan terintegrasi, memberikan wawasan dan analisis berharga yang tidak mungkin dilakukan dengan sistem pemrosesan data tradisional. Pada artikel ini, kita akan membahas prinsip-prinsip data warehouse dan memberikan studi kasus untuk menunjukkan penerapan praktisnya.

### 1. Data Integration

Data Integration adalah proses menggabungkan data dari berbagai sumber menjadi satu tampilan terpadu. Data warehouse bergantung pada prinsip ini



untuk memastikan bahwa semua data akurat, konsisten, dan mutakhir. Ini melibatkan penggunaan tools ETL (ekstrak, ubah, muat) untuk mengekstrak data dari sumber yang berbeda, mengubahnya menjadi format standar, dan memuatnya ke data warehouse.

**Studi Kasus:** Perusahaan ritel memiliki banyak sistem yang menyimpan informasi pelanggan, data penjualan, dan informasi inventaris. Dengan mengintegrasikan data ini ke dalam data warehouse, perusahaan memperoleh pandangan komprehensif tentang operasinya dan dapat membuat keputusan yang tepat untuk meningkatkan proses bisnisnya.

## 2. Historical Data

Data warehouse melibatkan penyimpanan data historis, yang memberikan pandangan lengkap tentang kinerja dari waktu ke waktu. Data historis berguna untuk analisis tren, peramalan, dan mengidentifikasi pola atau anomali yang dapat memengaruhi kinerja bisnis.

**Studi Kasus:** Lembaga keuangan melacak transaksi pelanggan, investasi, dan saldo rekening. Dengan menyimpan data ini di data warehouse, institusi memperoleh wawasan historis tentang perilaku pelanggan, yang dapat digunakan untuk mengembangkan kampanye pemasaran bertarget dan strategi investasi.

## 3. Non-Volatile Data

Data warehouse melibatkan penyimpanan data non-volatile, yang berarti bahwa data yang dimasukkan tidak dapat diubah, diperbarui, atau dihapus. Prinsip ini memastikan bahwa data akurat dan andal, yang penting untuk proses pengambilan keputusan.

**Studi Kasus:** Penyedia layanan kesehatan menyimpan catatan pasien, riwayat medis, dan rencana perawatan di data warehouse. Dengan memastikan bahwa data ini tidak mudah berubah, penyedia dapat memelihara catatan kesehatan pasien yang lengkap dan akurat, yang penting untuk memberikan perawatan yang efektif dan mengelola hasil pasien.

#### 4. Query Performance

Data warehouse melibatkan penggunaan tools untuk memastikan kinerja kueri yang cepat. Prinsip ini melibatkan pengoptimalan desain data warehouse dan strategi pengindeksan untuk memberikan waktu respons yang cepat untuk kueri yang kompleks.

**Studi Kasus:** Penyedia telekomunikasi menyimpan catatan detail panggilan dan data kinerja jaringan di data warehouse. Dengan mengoptimalkan desain data warehouse dan strategi pengindeksan, penyedia dapat dengan cepat menganalisis data kinerja jaringan, mengidentifikasi masalah, dan menyelesaikannya secara real-time.

#### 5. Data Access

Data warehouse melibatkan penyediaan akses yang aman dan terkontrol ke data. Prinsip ini memastikan bahwa data diakses hanya oleh personel yang berwenang dan bahwa akses dibatasi pada data dan fungsi yang diperlukan.

**Studi Kasus:** Instansi pemerintah menyimpan data sensitif terkait keamanan nasional di data warehouse. Dengan memberikan akses yang aman dan terkontrol ke data ini, agensi dapat memastikan bahwa hanya personel yang berwenang yang dapat mengaksesnya, sehingga mengurangi risiko pelanggaran keamanan.

## C. Metadata

Metadata didefinisikan sebagai data tentang data. Dalam konteks data warehouse, metadata memberikan informasi tentang data yang disimpan di data warehouse. Ini dapat mencakup informasi tentang sumber data, formatnya, strukturnya, dan kualitasnya. Metadata sangat penting dalam memungkinkan organisasi untuk memahami dan menggunakan data yang disimpan di data warehouse. Ini memberikan pemahaman yang jelas tentang data apa yang disimpan di data warehouse, dari mana asalnya, dan bagaimana data tersebut dapat digunakan untuk mendukung proses pengambilan keputusan. Berikut penjelasan kenapa metadata penting dalam suatu data warehouse:

### 1. Data Integration and Management

Metadata memainkan peran penting dalam integrasi dan manajemen data. Dengan memberikan informasi tentang data yang disimpan di data warehouse, metadata membantu memastikan bahwa data tersebut akurat, konsisten, dan terkini. Ini memungkinkan integrasi data di berbagai sumber dan memastikan bahwa data dikelola secara efektif.

### 2. Data Quality and Accuracy

Metadata membantu memastikan bahwa data yang disimpan di data warehouse memiliki kualitas dan akurasi yang tinggi. Ini memberikan informasi tentang struktur dan format data, memungkinkan organisasi untuk mengidentifikasi dan memperbaiki ketidakkonsistenan atau kesalahan dalam data.



### 3. Query Optimization

Metadata dapat digunakan untuk mengoptimalkan kinerja kueri. Dengan memberikan informasi tentang struktur data, metadata memungkinkan organisasi mengembangkan strategi kueri efektif yang meningkatkan waktu respons kueri dan memungkinkan pengambilan keputusan lebih cepat.

### 4. Governance and Compliance

Metadata memainkan peran penting dalam memastikan tata kelola dan kepatuhan terhadap peraturan dan kebijakan. Ini memberikan informasi tentang data yang disimpan di data warehouse, memungkinkan organisasi mengelola akses dan penggunaan data sesuai dengan peraturan dan kebijakan.

## **Studi Kasus: Pentingnya Metadata dalam Analisis Perusahaan Ritel**

Perusahaan ritel mengumpulkan data dari berbagai sumber, termasuk sistem titik penjualan, situs web e-commerce, dan platform media sosial. Perusahaan menyimpan data ini di data warehouse untuk mendukung proses pengambilan keputusannya. Namun, perusahaan kesulitan untuk memanfaatkan data secara efektif karena kurangnya pemahaman tentang data yang disimpan di gudang.

Untuk mengatasi masalah ini, perusahaan menerapkan sistem manajemen metadata. Sistem ini memberikan informasi tentang struktur dan format data yang disimpan di gudang, memungkinkan perusahaan untuk mengidentifikasi

dan memperbaiki ketidakkonsistenan atau kesalahan dalam data. Ini juga memberikan informasi tentang sumber data, memungkinkan perusahaan untuk melacak asal data dan menilai kualitasnya.

Sebagai hasil dari penerapan sistem manajemen metadata, perusahaan retail mampu meningkatkan keakuratan dan kualitas data yang disimpan di data warehouse. Hal ini memungkinkan perusahaan untuk mengembangkan kemampuan analitik dan pelaporan yang lebih efektif, yang menghasilkan pengambilan keputusan yang lebih baik dan peningkatan pendapatan.





## REFERENCE

<https://www.toptal.com/data-science/data-warehouse-concepts-principles>

<https://towardsdatascience.com/building-a-data-warehouse-basic-architectural-principles-66bd7059ffd0>

<https://dataschool.com/data-governance/single-source-of-truth/>

<https://www.javatpoint.com/what-is-meta-data>