Data Warehouse Design for Enterprise Business Intelligence

Abstract

Modern enterprises operate in highly digitized environments, generating massive volumes of structured and semi-structured data from diverse operational systems such as sales, finance, human resources, supply chain, and customer management platforms. Although these systems are optimized for routine transaction processing, they are not designed to support complex analytical queries, historical analysis, or strategic decision-making. As a result, organizations often face challenges such as fragmented data, inconsistent reports, poor query performance, and limited visibility into business trends. To overcome these limitations, this project focuses on the design and implementation of a centralized enterprise data warehouse to enable efficient Business Intelligence (BI) and advanced analytics.

The proposed solution adopts a dimensional modeling methodology, employing star and snowflake schemas to structure data into fact and dimension tables that support efficient Online Analytical Processing (OLAP). A comprehensive Extract, Transform, and Load (ETL) pipeline is developed to integrate data from multiple heterogeneous source systems. The ETL process includes data extraction, cleansing, validation, transformation, and aggregation to ensure high data quality. Common data issues such as missing values, redundancy, inconsistency, and format mismatches are systematically resolved to provide a single, reliable version of enterprise data.

On top of the data warehouse, BI tools are utilized to design interactive dashboards and analytical reports that visualize key performance indicators (KPIs), trends, and patterns across business functions. The system is evaluated using metrics such as query response time, scalability, data accuracy, and reporting efficiency. Experimental results show significant improvements in analytical performance and data consistency compared to traditional operational databases. Overall, the proposed data warehouse framework provides a scalable, extensible, and reliable foundation for enterprise-level business intelligence, supporting data-driven decision-making and long-term strategic planning.