DeivaKumaran Dhanasegaran
Aswathnarayan
Pranesh Ramakrishnan

Database Design and Data Management - Fall 2016

Final Project – Modelling IMDB database

Project Summary:

Designed and developed a database for IMDB Movie Database System using SQL, Oracle and Hive database.IMDB, the world's most popular and authoritative source for movie, TV and celebrity content. In this Project we followed Industrial Standard. First we gathered Business Requirement from IMDB web site and Developed Conceptual Data Model followed by Logical Data Model (Normalization is the last part of the logical design).

In the Logical data Model we normalized the database to reduce data redundancy and to ensure data integrity in which Insert, delete and Update anomalies is minimized.

In physical data model finally deployed the application in SQL and Oracle Server. In Physical Data Model we have designed Entity Relationship Diagram for SQL and Oracle and reverse engineered, generated DDL/SQL scripts and alter scripts using Toad data Modeler.

Used Toad data Modeler to design Entity Relationship Diagrams of particular database platform, convert logical data model to physical model, load ER Diagram directly from your database (Reverse Engineering feature, loading of DDL scripts), update physical models, generate DDL/SQL scripts and also alter scripts.Finally implemented and deployed the user end functionalities using Procedures, Triggers, Views, Back-ups and Sub-queries.

Deployed the IMDB database in Apache Hive, data warehouse infrastructure built on top of Hadoop for providing data summarization, query, and analysis for large amount of data.

We used Tableau which is Business Intelligence tool for visually analyzing the data and for Reporting the trends, variations and density of the data in form of graphs and charts. Generated statistics of top 10 movies according to years, most revenue generated movie of particular year, most rated movies by genre, Most liked genre by country based on revenue, etc.