# **REPORT PROVIDE BY: YAN MARCHAN Music Streaming Service: E-commerce System Development Report**

## **1. Introduction**

This report provides a comprehensive overview of constructing an e-commerce application for a Music Streaming Service. The project includes the development of both Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP) databases, data migration from CSV files, and the implementation of an ETL process for data transferring and reporting.

## **2. OLTP System Development**

### **2.1 ER-Diagram Overview**

The ER-Diagram for the music streaming service depicts various entities like Users, Artists, Albums, Tracks, Playlists, and Subscription Plans, and their relationships. This diagram is fundamental in understanding the relational structure and is pivotal for database design. The diagram can be viewed [here](https://media.geeksforgeeks.org/wp-content/uploads/20240226185635/MusicStreamingER-(1).webp).

### **2.2 Database Schema Creation**

The OLTP database schema includes several tables designed to manage the daily operations of the music streaming service:

#### **Tables:**

* **User**: Stores user details including username, email, and password.
* **Artist**: Contains artist details like name and genre.
* **Album**: Records details of albums including title, associated artist, and release date.
* **Track**: Maintains track information such as title, associated artist, album, play count, duration, and release date.
* **Playlist**: Holds information about user-created playlists.
* **Like**: Manages user likes for tracks.
* **SubscriptionPlan**: Details about different subscription plans available.
* **Payment**: Tracks payment transactions made by users for subscriptions.

### **2.3 Functions and Procedures**

Several SQL functions and stored procedures are implemented to support operations such as adding/removing likes, processing payments, and managing albums and artists.

## **3. OLAP System Development**

### **3.1 Design of Dimension and Fact Tables**

The OLAP system is structured to facilitate complex queries and generate reports efficiently.

#### **Dimension Tables:**

* **dim\_user**: Includes user attributes along with subscription start and end dates.
* **dim\_artist**: Stores artist details.
* **dim\_album**: Contains album information with release dates.
* **dim\_track**: Tracks details along with play counts and duration.
* **dim\_time**: Time dimensions like date, month, quarter, and year.

#### **Fact Tables:**

* **fact\_streaming**: Records streaming data by track, including play counts and listening time.
* **fact\_subscription**: Tracks subscription data by user, including the start date, end date, and monthly fee.

### **3.2 Indexes**

Indexes are created on the fact tables to enhance query performance, particularly focusing on track ID and subscription data.

## **4. Data Loading and ETL Process**

### **4.1 Data Loading from CSV Files**

Data is initially loaded into the OLTP system from CSV files using PostgreSQL's COPY command. This includes data for users, artists, albums, tracks, and more, ensuring that duplicate entries are avoided.

### **4.2 ETL Implementation**

An ETL process, implemented in Python using pandas and SQLAlchemy, extracts data from the OLTP database, transforms it (e.g., calculating listening times, setting current subscription statuses), and loads it into the OLAP database. This process ensures that the data in the OLAP database is up-to-date and ready for analysis.

### **4.3 CSV Data Generation**

Scripts using the Faker library generate realistic dummy data for initial testing, covering users, artists, albums, tracks, and subscription plans.

## **5. Conclusion**

The development of the Music Streaming Service's e-commerce system incorporates robust database design and efficient data handling mechanisms to support both transactional and analytical operations. The system is designed to handle large volumes of data efficiently, ensuring scalability and reliability. Through careful planning and execution of the OLTP and OLAP systems, alongside a comprehensive ETL process, the platform is well-equipped to provide a seamless user experience and insightful data analysis.