Advanced Data Management - DVD Rental Business Analysis

Overview

In this project, I developed an automated reporting solution for a DVD rental business using a PostgreSQL database. The database includes key information about rentals, films, categories, and inventory.

Sample database source: The **dvdrental.zip** file containing dvdrental database can be downloaded from **neon.tech** here: dvdrental.zip.

Technologies used:
- **PostgreSQL** for database management and SQL queries.
- **pgAdmin** as the database administration tool for interacting with PostgreSQL.
- **PL/pgSQL** for custom functions, triggers, and stored procedures.

Key Tasks

Business Report Creation:
- Defined key business questions (e.g., rental patterns and category popularity) based on the dataset.
- Created two key tables for the report:
  - **Detailed Table:** Shows individual rental transactions with rental ID, year and month of rental, and the film category.
  - **Summary Table:** Provides aggregated rental counts per film category for each month and year.

Data Transformation & Advanced SQL Queries:
- Created **user-defined functions** to extract the year and month from the rental_date field in the rental table. This transformed data was used to populate the year_of_rental and month_of_rental columns in the detailed table.
- Wrote complex SQL queries that utilized multiple joins in a single query, combining data from several tables (rental, inventory, film_category, and category).
- Implemented a **trigger** to automatically update the summary table whenever new data is added to the detailed table, ensuring the report remains up-to-date in real-time.
- Developed a **stored procedure** to refresh both the detailed and summary tables, clearing old data and performing new data extraction and aggregation.

Project Deliverables

SQL script: the "dvdrental_project.sql" file contains several key components:
- **Data transformation functions**: It includes functions to extract the month and year from rental dates, enabling time-based reporting and analysis.
- **Table creation**: The script creates two main tables—detailed_table (for storing rental and category data) and summary_table (for aggregated rental statistics).
- **Data extraction and population**: It extracts raw data from the source tables and populates the detailed_table with relevant rental and category information.

- **Automated updates with triggers**: A trigger updates the summary_table whenever new data is inserted into the detailed_table, ensuring that summary data is always up-to-date.
- **Data refresh procedure**: The refresh_tables procedure allows for refreshing both the detailed_table and summary_table with the latest data, providing a clean and accurate dataset for reporting.
- **Reset the environment (optional)**: The part includes code to drop all the created functions, triggers, and tables, which is useful for resetting the environment.

Result tables:
- "detailed_table.csv": Contains the detailed rental data with the year, month, and category of each rental.
- "summary_table.csv": Contains the summary data showing the total number of rentals for each film category, organized by year and month.

Business Impact

This automated reporting solution enables the DVD rental store to track rental trends over time, identify popular film categories, and make data-driven decisions for inventory and marketing. By having up-to-date rental statistics available, the business can better forecast demand, optimize inventory levels, and tailor promotions to high-performing film categories.