Rajeshwar Shinde
12th July 2023
Kaggle Dataset Link: https://rb.gy/imo0m

Music Store Dataset Project Documentation

Introduction

This project aims to analyse a music store dataset and gain insights into various aspects of the store's operations. The dataset includes information about artists, albums, tracks, customers, invoices, employees, playlists, and more. The analysis is performed using PostgreSQL and PgAdmin 4 v6.

Dataset Overview

The dataset consists of the following tables:

1. Artist

- ArtistId (primary key)
- Name

2. Playlist

- PlaylistId (primary key)
- Name

3. Employee

- EmployeeId (primary key)
- LastName
- FirstName
- Title
- ReportsTo
- BirthDate
- HireDate
- Address

- City
- State
- Country
- PostalCode
- Phone
- Fax
- Email

4. Album

- AlbumId (primary key)
- Title
- ArtistId (foreign key referencing Artist table)

5. PlaylistTrack

- PlaylistId (foreign key referencing Playlist table)
- TrackId (foreign key referencing Track table)

6. Customer

- CustomerId (primary key)
- FirstName
- LastName
- Company
- Address
- City
- State
- Country
- PostalCode
- Phone
- Fax
- Email
- SupportRepId (foreign key referencing Employee table)

7. Track

- TrackId (primary key)
- Name
- AlbumId (foreign key referencing Album table)

- MediaTypeId (foreign key referencing MediaType table)
- GenreId (foreign key referencing Genre table)
- Composer
- Milliseconds
- Bytes
- UnitPrice

8. InvoiceLine

- InvoiceLineId (primary key)
- InvoiceId (foreign key referencing Invoice table)
- TrackId (foreign key referencing Track table)
- UnitPrice
- Quantity

9. Invoice

- InvoiceId (primary key)
- CustomerId (foreign key referencing Customer table)
- InvoiceDate
- BillingAddress
- BillingCity
- BillingState
- BillingCountry
- BillingPostalCode
- Total

10. MediaType

- MediaTypeId (primary key)
- Name

11. Genre

- GenreId (primary key)
- Name

Tools Used

- PgAdmin 4 v6: The graphical administration tool used for managing the PostgreSQL database and executing queries.

Project Workflow

1. **Dataset Acquisition**: The music store dataset was obtained from Kaggle.
2. **Database Setup**: The PostgreSQL database was set up using PgAdmin 4 v6.
3. **Data Import**: The dataset was imported into the PostgreSQL database using appropriate data import techniques.
4. **Database Design**: The database schema was designed based on the provided dataset and its relationships.
5. **Data Analysis**: Various SQL queries and operations were performed on the database to extract insights and answer specific questions.
6. **Data Visualization**: Power BI was utilized to create visual representations of the data analysis results. The visualizations included cards, bar graphs, donut charts, pie diagrams, year-wise filter tiles, maps, and column charts to showcase key metrics, sales trends, customer behaviour, and geographical insights.
7. **Results and Documentation**: The analysis results were recorded in an Excel file, with each question's answer presented in a separate sheet.
8. **Project Documentation**: This document summarizes the project, including dataset details, tools used, and the project workflow.

SQL Analysis Results

The analysis of the music store dataset yielded insights into various aspects of the store's operations. The answers to specific questions can be found in the accompanying Excel file. The insights gained include:

- Sales performance analysis based on invoices and customers.
- Employee hierarchy
- Popular genres, tracks, and albums.
- Impact of specific artists and their tracks on total sales
- Artist information and their associations with albums.
- Media types and genres of tracks.

Power BI Report Analysis

The report, titled "Sales Analysis," provides an in-depth analysis of sales performance in the music store. It includes a range of visualizations and key metrics to present a comprehensive view of sales trends and customer behaviour.

The report utilizes various visualizations, such as bar graphs, donut charts, pie diagrams, year-wise filter tiles, maps, and column charts, to effectively showcase different aspects of the sales analysis.

Key visualizations and features of the report include:

- **Cards for Key Metrics**: The report includes cards to display key metrics, such as total sales revenue, number of customers, total number of tracks, number of invoices, and the total artist & albums. These cards provide a quick snapshot of important performance indicators.
- **Bar Graphs**: Bar graphs are used to present comparative sales data across different dimensions. For example, they may show sales by genre, artist or albums allowing users to analyse sales performance across various categories.
- **Donut Chart and Pie Diagram**: Donut charts and pie diagrams offer a visual representation of sales distribution by different criteria. They can display the percentage of sales contributed by each sales executive or count of customers handled by a specific sales executive.
- **Year-wise Filter Tiles**: The report includes interactive year-wise filter tiles, allowing users to select a specific year to analyse sales trends and performance within that period. This feature enables users to dynamically explore data and perform year-to-year comparisons.
- **Map Visualization**: A map visualization is utilized to showcase country-wise sales data. By showing different sized bubbles on countries based on sales revenue, the map provides a geographical perspective on sales distribution, allowing users to identify high-performing regions and potential growth opportunities.
- **Column Chart for Sales Trends**: A column chart is used to present sales trends over time, allowing users to analyse sales performance across different years and quarters. This visualization can highlight seasonal variations and identify periods of significant growth or decline.

The combination of these visualizations and features in the "Sales Analysis" report provides a comprehensive understanding of sales performance, customer behaviour, and market trends within the music store dataset.

Conclusion

The "Music Store Dataset" project involved analysing a comprehensive dataset using PostgreSQL and PgAdmin 4 v6. Through SQL queries and operations, valuable insights were gained regarding the music store's sales, employees, customers, artists, playlists, and tracks. The project documentation, along with the Excel file containing the analysis results, provides a comprehensive overview of the project and its outcomes.