Overview

This project is designed for beginners to learn how to analyze a music database using SQL. By exploring the dataset, users will gain insights into the store's business growth, customer behavior, and music preferences. The analysis will help the store understand trends and make data-driven decisions.

Objectives

This project will help you:

- Explore and understand the music store's sales data.
- Use SQL queries to retrieve, filter, and aggregate information.
- Answer important business questions related to customer purchases, popular genres, top artists, and sales trends.
- Gain insights into the store's growth and customer preferences based on data-driven analysis.

Scope

In-Scope

➔ Data extraction from a relational database.
➔ SQL queries for basic and advanced analysis.
➔ Identifying top genres, artists, and customers.
➔ Analyzing sales trends.
➔ Insights into regional customer behavior.

Out of Scope

★ Machine Learning-based music recommendations.
★ Predictive analytics beyond SQL.
★ Web development or UI/UX design for visualization.

Tools and Technologies

➢ Database Management System: Oracle SQL Developer
➢ Query Language: SQL

Dataset Description

The dataset consists of multiple tables, including:

Table Name	Description
music_employee	Contains details about employees such as name, role, and reports to.
music_customer	Contains customer details such as name, email, and country.
music_invoice	Stores purchase transaction details.
music_invoice_line	Line-by-line breakdown of purchases.
music_track	Details about each track including genre and album.
music_album	Stores album information.
music_artist	Contains artist details.
music_genre	Lists different music genres.

Queries (Solution of all queries: check the *music_data_analysis.sql* file)

Queries (Set-1: Easy):

Query-1: Who is the senior most employee based on job title?

Query-2: Find the list of youngest employees based on job title?

Query-3: Find the details that have lived in Lethargies.

Query-4: Which countries have the most invoices?

Query-5: What are the top three values of total invoice?

Query-6: Which city has the best customers? We would like to throw a promotional music festival in the city where we made the most money. Write a SQL query that returns one city that has the highest sum of invoice totals. Return both the city name and sum of all invoice totals.

Query-7: Who is the best customer? The customer who has spent the most money will be declared the best customer. Write a SQL query that returns the person who has spent the most money.

Query-8: Find the last 5 customers? The customer who has spent the least money will be declared the last customer. Write a SQL query that returns the person who has spent the least money with the city name also.

Query-9: Find the list of customers id, customers full name, billing city and billing country with the total invoice is greater than 100.

Query-10: Calculate the total invoice based on the billing city.

Queries (Set-2: Intermediate):

Query-11: Write a SQL query to return the email, first name, last name and genre of all 'Rock' music listeners. Return your list ordered alphabetically by email starting with A.

Query-12: Let's invite the artists who have written the most rock music in our dataset. Write a SQL query that returns the artist name and total track count of the top 10 rock bands.

Query-13: Return all the track names that have a song length longer than the average song length. Return the name and milliseconds for each track, ordered by the song length with the longest songs listed first.

Query-14: Write a SQL query to return all the track ids, track names, milliseconds and percentage for each track, ordered by the highest percentage.

Queries (Set-3: Advanced):

Query-15: Find how much amount spent by each customer on artists? Write a SQL query to return customer name, artist name and total spent.

Query-16: We want to find out the most popular music genre for each country. We determine the most popular genre as the genre with the highest amount of purchases. Write a SQL query that returns each country along with the top genre. For countries where the maximum number of purchases is shared, return all genres.

Query-17: Write a SQL query that determines the customer that has spent the most on music for each country. Write a SQL query that returns the country along with the top customer and how much they spent. For countries where the top amount spent is shared, provide all customers who spent this amount.

Queries (Set-4: Mixed):

Query-18: What are the most popular genres of music?

Query-19: What are the 5 most popular artists?

Query-20: What are the 5 most popular songs?

Query-21: What are the average prices of different types of music?

Query-22: What are the most popular countries for music purchases?

Expected Outcomes

- Ability to analyze music sales and customer behavior using SQL.
- Understanding of how to extract business insights from a database.
- Improved SQL query writing skills.
- Identification of top-performing genres, artists, and customers.
- Clear visualization of store growth and revenue trends.

Future Works

- Data visualization: Use Python (Matplotlib/Seaborn) or Power BI.
- Advanced SQL queries: Implement window functions and CTEs.
- Predictive analytics: Forecast sales trends using machine learning.

Project Completion Milestones

★ Understanding database schema
★ Writing and executing SQL queries
★ Analyzing and interpreting query results
★ Documenting key business insights
★ Exploring future data enhancements

Conclusion

This project provides a hands-on approach to SQL-based music data analysis. By working with a structured music playlist database, beginners will learn how to extract insights that can drive business decisions. Through this project, users will develop essential data analysis skills applicable in real-world data science and analytics careers.

End of Document