**Objective Questions**

1. Does any table have missing values or duplicates? If yes how would you handle it ?

**Approach**

* To identify missing values, I conducted a null check across relevant columns in key tables (employee, customer, and track) using SQL queries with WHERE and OR conditions to capture nulls in specific fields.

**Explanation**

Employee Table:

* Null values were checked for key column reports\_to.
* To handle nulls, the COALESCE function was applied to substitute missing values in reports\_to with "N/A".

Customer Table:

* Null values were identified in columns like company, state, postal\_code, and phone.
* We handled these by using COALESCE to provide suitable defaults, e.g., "N/A" for company, "Unknown" for state, and "Not Provided" for phone.

Track Table:

* Columns like composer were checked, and missing values were replaced with "N/A" using COALESCE.

**Result**

* All other tables were reviewed, and no missing values or duplicates were found.
* The COALESCE function successfully replaced null values in each table with appropriate default values, preserving data usability without modifying the original records.
* No duplicate rows were found across primary key columns in any table.

**Code**

A screenshot of a computer program

Description automatically generated

A screenshot of a computer screen

Description automatically generated

A white sheet with black text

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a spreadsheet

Description automatically generated

1. Find the top-selling tracks and top artist in the USA and identify their most famous genres.

**Approach**

* Utilized SQL joins and grouping to find top-selling tracks, the most popular artist, and their most famous genres in the USA. Queries were grouped by track name, artist, and genre, filtering for customers in the USA.

**Explanation**

* Top-Selling Tracks: Joined track, invoice\_line, invoice, and customer tables, filtered for USA, and grouped by track to calculate sales, ordered by quantity sold (top 15).
* Top Artist: Joined track, album, artist, invoice\_line, invoice, and customer tables, summing quantity sold per artist, ordered to find the top artist in the USA.
* Famous Genres: Used a subquery to locate the top artist, then joined genre, track, invoice\_line, and album tables to group and order genres by sales.

**Visualisation**

**Result**

* Top-Selling Tracks: Found the 15 most popular tracks in the USA based on sales, reflecting customer preferences.
* Top Artist: Identified Van Halen as the top-selling artist in the USA with 43 total sales, indicating high market demand.
* Popular Genres: Listed Rock as the most popular genre associated with Van Halen, with 43 total sales, revealing the favored music style among their fans.

**Code**

**A computer screen shot of a program code

Description automatically generatedA screenshot of a music list

Description automatically generated**

**A computer screen shot of a program

Description automatically generated**

**A screen shot of a computer

Description automatically generated**

****

1. What is the customer demographic breakdown (age, gender, location) of Chinook's customer base?

**Approach**

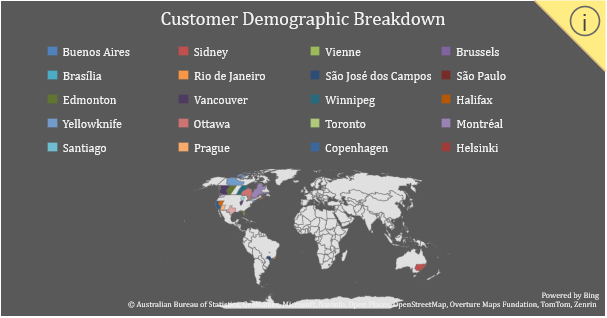
* I used SQL to analyze the customer demographic breakdown by aggregating data from the customer table, focusing on country, state, and city to capture the total number of customers in each demographic segment.

**Explanation**

Customer Demographic Breakdown:

* The query selected key columns: country, state, and city, counting the total number of customers (customer\_id) for each unique combination.
* I used the COALESCE function to replace any null values in the state column with "Not Available" for clarity.
* The results were grouped by country, state, and city, and ordered accordingly to present a clear demographic structure.

**Visualization**



**Result**

* The analysis provided a comprehensive breakdown of Chinook's customer demographics, detailing the total number of customers by country, state, and city, which can inform marketing strategies and customer engagement efforts.

**Code**

**A computer screen shot of a program code

Description automatically generatedA table with numbers and names

Description automatically generated**

1. Calculate the total revenue and number of invoices for each country, state, and city:

**Approach**

* I employed SQL to calculate the total revenue and the number of invoices for each combination of country, state, and city by joining the customer and invoice tables.

**Explanation**

Revenue and Invoice Calculation:

* The query selected country, state, and city, summing the total revenue (i.total) and counting the number of invoices (i.invoice\_id) for each grouping.
* The COALESCE function was used to substitute null values in the state column with "Not Available" for clarity.
* Results were grouped by country, state, and city, with a descending order based on total revenue to highlight the highest earning locations.

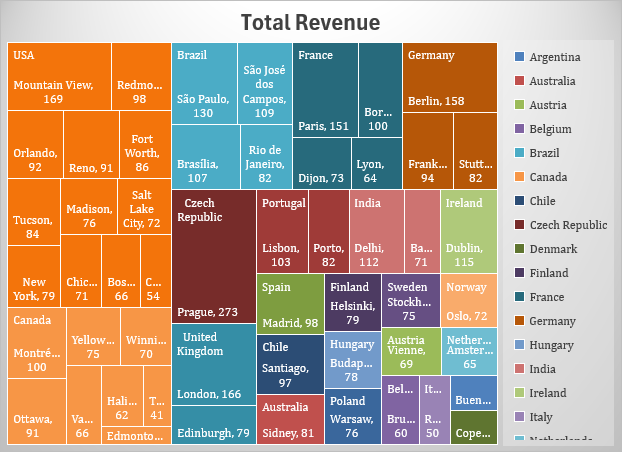
**Code**

**A computer screen with white text

Description automatically generatedA screenshot of a data

Description automatically generated**

**Visualization**



**Result**

* The analysis provided a detailed view of total revenue and the number of invoices for each country, state, and city, enabling insights into financial performance across different regions.

1. Find the top 5 customers by total revenue in each country

**Approach**

* I utilized Common Table Expressions (CTEs) in SQL to calculate total revenue for each customer and then ranked them within their respective countries to identify the top 5 customers by revenue.

**Explanation**

Revenue Calculation:

* The first CTE, country\_wise\_revenue, aggregates total revenue for each customer by joining the customer and invoice tables. It groups the results by customer ID and country, summing the i.total for total revenue calculation.

Ranking Customers:

* The second CTE, top\_customers, ranks the customers within each country based on their total revenue using the RANK() window function. This allows us to categorize customers by their revenue performance within their country.

Final Selection:

* The main query selects the customer ID, name, country, and total revenue from the ranked results, filtering to include only those customers whose ranking is 5 or lower.

**Visualization**

**Code**

**A screenshot of a computer program

Description automatically generatedA screenshot of a computer

Description automatically generated**

**Result**

* This analysis provided a comprehensive list of the top 5 customers by total revenue in each country, facilitating insights into customer value and sales performance across different regions.

1. Identify the top-selling track for each customer

**Approach**

* I used Common Table Expressions (CTEs) in SQL to first calculate the total quantity of tracks purchased by each customer and then determined the top-selling track for each customer based on these totals.

**Explanation**

Total Track Purchases:

* The first CTE, total\_track\_customer, aggregates the total quantity of tracks purchased by each customer by joining the customer, invoice, and invoice\_line tables. The results are grouped by customer ID and name, yielding the total quantity of tracks purchased by each customer.

Identifying Top Track:

* The second CTE, top\_track\_customer, ranks the tracks purchased by each customer based on total quantity sold. This is accomplished using the ROW\_NUMBER() window function, partitioning the results by customer ID and ordering by total quantity in descending order.

Final Selection:

* The main query retrieves the customer ID, customer name, top track name, and the total quantity of that top track, filtering to include only those entries where top\_rank is 1, ensuring that only the top-selling track for each customer is returned.

**Code**

**A screenshot of a computer program

Description automatically generatedA screenshot of a computer

Description automatically generated**

**Visualization**

**Result**

* This analysis produced a list of each customer along with their top-selling track, reflecting the most popular purchases for each individual customer.

1. Are there any patterns or trends in customer purchasing behavior (e.g., frequency of purchases, preferred payment methods, average order value)?

**Approach**

* I employed SQL queries to analyze various aspects of customer purchasing behavior, focusing on purchase frequency, average order value, and total revenue generated by each customer.

**Explanation**

Frequency of Purchases:

* The first query aggregates the number of purchases made by each customer per year. By joining the customer and invoice tables, I counted the number of invoices per customer for each year, helping identify any trends in purchasing frequency over time.

Average Order Value:

* The second query calculates the average order value for each customer by averaging the total amount from their invoices. This is done by joining the customer and invoice tables and grouping the results by customer ID to reflect individual spending behavior.

Total Revenue:

* The third query sums the total revenue generated by each customer by aggregating the total from their invoices. This provides insights into overall customer value and contribution to revenue, allowing for comparison among customers.

**Visualization**

**Code**

**A computer screen shot of a program

Description automatically generatedA screenshot of a table

Description automatically generated**

**A computer screen shot of white text

Description automatically generatedA screenshot of a computer

Description automatically generated**

**A computer screen with white text

Description automatically generatedA screenshot of a computer

Description automatically generated**

**Result**

* The analysis reveals patterns in customer purchasing behavior, such as the frequency of purchases, average spending per order, and total revenue per customer, which can inform marketing strategies and customer engagement efforts.

1. What is the customer churn rate?

**Approach**

* I used SQL queries to determine the customer churn rate by identifying customers who have not made a purchase in the past year.
* This involved creating subqueries to find the recent invoice date and checking for customers whose last purchase falls outside the defined timeframe.

**Explanation**

Churn Checking Date:

* The first subquery calculates the date one year prior to the most recent purchase date recorded in the invoice table. This serves as the threshold for determining customer churn.

Churn Customers:

* The second subquery identifies customers who have not made any purchases in the past year. This is achieved by joining the customer and invoice tables to find the last purchase date for each customer.
* Customers are flagged if their last purchase is either null (no purchases ever) or older than the churn checking date.

Churn Rate Calculation:

* The final query calculates the churn rate by dividing the number of churned customers by the total number of customers, then multiplying by 100 to express it as a percentage.

**Code**

**A computer screen shot of a program

Description automatically generated**

**Result**

* The calculated churn rate is 1.69%, indicating that a small proportion of customers have ceased making purchases over the past year. This metric is critical for evaluating business health and guiding retention strategies.

1. Calculate the percentage of total sales contributed by each genre in the USA and identify the best-selling genres and artists.

**Approach**

* I used SQL queries to calculate the total sales contributed by each genre in the USA, followed by determining the best-selling genres and artists. This involved joining multiple tables and aggregating sales data based on genre and artist.

**Explanation**

Genre-wise Sales Contribution:

* First Subquery: Calculates total sales for each genre by summing the quantity sold multiplied by the unit price. This data is filtered to include only customers from the USA.
* Total Sales Calculation: A second subquery aggregates total sales across all genres to find the overall sales figure.

Percentage Contribution:

* The final output displays each genre's name, total sales, and the percentage contribution to overall sales in the USA.

Best-Selling Genres and Artists:

* Sales Calculation by Genre and Artist: Another subquery sums sales while grouping by genre and artist names to identify which artists are driving sales in each genre.
* Ranking Artists: The final selection ranks artists within each genre based on sales, highlighting the top performers.

**Visualization**

**Code**

**A screenshot of a computer screen

Description automatically generatedA table of music sales

Description automatically generated with medium confidence**

**A computer screen shot of a black background

Description automatically generatedA screenshot of a music list

Description automatically generated**

**Result**

* The query provides a detailed breakdown of each genre's sales contribution in the USA and identifies the best-selling genres alongside their leading artists.

1. Find customers who have purchased tracks from at least 3 different genres

**Approach**

* The query identifies customers who have bought tracks from three or more distinct genres by using SQL joins and aggregation functions.

**Explanation**

* Table Joins: The query joins the customer, invoice, invoice\_line, track, and genre tables to link customers to their purchases and the genres of those tracks.
* Counting Genres: It uses COUNT(DISTINCT g.genre\_id) to count the unique genres associated with each customer.
* Filtering: The HAVING clause filters the results to include only those customers with a genre count of three or more.

**Visualization**

**Code**

**A computer screen with text

Description automatically generatedA screenshot of a computer

Description automatically generated**

**Result**

* The output includes customer IDs, names, and the count of distinct genres they've purchased from, ordered by genre count in descending order.

1. Rank genres based on their sales performance in the USA

**Approach**

* The query ranks music genres in the USA based on total sales revenue using SQL aggregation and ranking functions.

**Explanation**

* Table Joins: It joins the track, invoice\_line, invoice, customer, and genre tables to gather information on sales by genre.
* Calculating Sales: The query calculates total sales performance by summing the product of quantity sold and the unit price of each track.
* Grouping and Ordering: Results are grouped by genre name and ordered by sales performance in descending order.
* Ranking: The DENSE\_RANK() function assigns a rank to each genre based on its sales performance.

**Visualization**

**Code**

**A screen shot of a computer program

Description automatically generatedA table with numbers and text

Description automatically generated**

**Result**

* The output displays genre names, their sales performance figures, and their rankings, providing insights into which genres are the most successful in terms of sales in the USA.

1. Identify customers who have not made a purchase in the last 3 months

**Approach**

* The query identifies customers who have not made any purchases within the last three months by checking their invoice dates against the current date.

**Explanation**

* Table Join: It joins the customer and invoice tables to link customers with their purchase records.
* Filtering: The query filters invoices to include only those with dates older than three months from the current date using CURDATE() - INTERVAL 3 MONTH.
* Selecting Distinct Customers: It selects distinct customer IDs and their names to avoid duplicates in the results.

**Code**

**A computer code with white text

Description automatically generatedA screenshot of a computer

Description automatically generated**

**Result**

* The output lists the customer IDs and names of customers who have not made any purchases in the last three months, aiding in customer retention strategies.

**Subjective Questions**

1. Recommend the three albums from the new record label that should be prioritised for advertising and promotion in the USA based on genre sales analysis.

**Approach**

* Identify the top-performing albums from a new record label in the USA based on total sales, categorized by genre.

**Explanation**

* Sales Calculation: The query calculates total sales for each album by genre using sum(il.quantity \* t.unit\_price), representing sales analysis.
* Genre and Album Grouping: It groups the results by genre\_name and album title (al.title) to evaluate performance within each genre.
* Ranking: A DENSE\_RANK() window function is applied to order albums based on their total sales, with the highest sales given the highest ranking.
* Filter for USA: The WHERE c.country = "USA" filter limits the analysis to the USA customer base.
* Limit Top 3 Results: It limits the output to the top three albums by ranking to provide clear recommendations.

**Code**

**A computer screen shot of a program

Description automatically generatedA screenshot of a computer

Description automatically generated**

**Insights**

* The top three rock albums, From The Muddy Banks Of The Wishkah [Live] and Are You Experienced? (both at 27.72 sales units), along with The Doors (26.73 sales units), demonstrate strong sales performance and popularity. These albums should be prioritized for advertising and promotion to maximize reach and capitalize on their existing success in the market.

**Result**

* The output will list the top three albums from the new record label, their genres, and total sales in the USA, along with their rank. This information can be used to guide promotion efforts for the albums most likely to yield high returns.

1. Determine the top-selling genres in countries other than the USA and identify any commonalities or differences.

**Approach**

* Calculate total sales for each genre separately in non-USA and USA regions.
* Combine these results to assess genre popularity globally.
* Calculate and compare percentage contributions by genre within each region.

**Explanation**

* We first determine genre sales by totaling quantities in non-USA countries, followed by a similar calculation within the USA. By joining these results, we obtain a comprehensive view of genre popularity, allowing us to calculate the percentage each genre contributes in each region.

**Visualization**

**Insights**

* TV Show Genre: The total percentage of sales for the TV Show genre is almost equally divided between the USA and other countries, indicating a balanced global appeal.
* Hip Hop/Rap Genre: This genre is more popular in the USA, with 60% of its sales coming from the USA and 40% from other countries, showing a stronger affinity among American audiences.
* Classical Genre: Classical music is significantly more popular outside the USA, with 90% of its sales from other countries and only 10% from the USA, suggesting a preference for this genre internationally over domestically.

**Code**

A screen shot of a computer code

Description automatically generatedA screenshot of a music list

Description automatically generatedA screenshot of a computer program

Description automatically generatedA screenshot of a data

Description automatically generated

**Result**

* The output ranks genres by global sales and shows the percentage split between USA and non-USA regions. This reveals which genres are universally popular and which appeal more strongly to specific regions, informing data-driven marketing and promotional strategies.

1. Customer Purchasing Behavior Analysis: How do the purchasing habits (frequency, basket size, spending amount) of long-term customers differ from those of new customers? What insights can these patterns provide about customer loyalty and retention strategies?

**Approach**

* Customer Segmentation: Classify customers into 'New' and 'Long-Term' based on purchase history.
* Metrics Calculation: Compute average purchase frequency, total items, and spending for each segment.

**Explanation**

* The analysis compares purchasing habits between new and long-term customers to identify loyalty patterns and inform retention strategies.

**Insights**

* Long-term customers have higher purchase frequency and spending.
* Average Purchase Frequency: 80.63
* Average Basket Size: 80.63
* Average Spending Amount: 805.14
* Average Order Value: 9.77

**Code**

**A screenshot of a computer program

Description automatically generated**

**Result**

* This analysis provides actionable insights for enhancing customer loyalty and retention strategies.

1. Product Affinity Analysis: Which music genres, artists, or albums are frequently purchased together by customers? How can this information guide product recommendations and cross-selling initiatives?

**Approach**

* Affinity Analysis: Identify combinations of genres, artists, and albums frequently purchased together by leveraging invoice data.
* Data Grouping: For each analysis, group the data by combinations and count instances of co-purchases to determine affinity strength.

**Explanation**

* This analysis explores customer purchasing behavior to uncover relationships between different music genres, artists, and albums. Understanding which products are often bought together can enhance product recommendations and inform cross-selling strategies.

**Code**

**A screenshot of a computer program

Description automatically generatedA screenshot of a music list

Description automatically generatedA screenshot of a computer program

Description automatically generatedA screenshot of a music list

Description automatically generatedA screenshot of a computer program

Description automatically generatedA screenshot of a music album

Description automatically generated**

**Insights**

* Genre Affinity: Identify genres that customers often purchase together, allowing targeted marketing strategies for similar genres.
* Artist Affinity: Recognize pairs of artists frequently bought together, suggesting potential collaborations or bundled offerings.
* Album Affinity: Highlight albums that are popular among similar customer groups, aiding in crafting personalized recommendations.

**Result**

* These findings can lead to more effective marketing strategies, enhancing customer experience and increasing sales through targeted recommendations.

1. Regional Market Analysis: Do customer purchasing behaviors and churn rates vary across different geographic regions or store locations? How might these correlate with local demographic or economic factors?

**Approach**

* Purchase Frequency and Spending Analysis: Calculate each customer's purchase frequency, total spending, and average order value, then group this data by region (country, state, city).
* Churn Rate Calculation: Identify churned customers who haven’t purchased within the past year, grouped by region to determine churn rates.

**Explanation**

* This query assesses customer purchasing behavior and churn rates by geographic region.
* The purchase\_frequency subquery gathers each customer’s spending and frequency data, while customer\_region segments this data by location, showing spending trends and average order values for each area.
* The second part identifies churned customers by region based on their last purchase date.

**Code**

**A screen shot of a computer code

Description automatically generatedA table with numbers and a list of flights

Description automatically generated with medium confidenceA screen shot of a computer code

Description automatically generatedA table of numbers with names

Description automatically generated with medium confidence**

**Result**

The query provides:

* Total customers, total purchases, and total spending for each region.
* Average order value, indicating purchasing trends in each location.
* Churn rate by region, showing customer retention challenges across locations.
* This analysis helps in understanding regional differences in customer engagement and spending, supporting targeted strategies for marketing and retention

1. Customer Risk Profiling: Based on customer profiles (age, gender, location, purchase history), which customer segments are more likely to churn or pose a higher risk of reduced spending? What factors contribute to this risk?

**Approach**

* Customer Profiling: Categorize customers based on their last purchase date, total spending, and average order value, grouping them into "High Risk," "Medium Risk," or "Low Risk."
* Risk Summary by Region: Summarize data by region and risk profile, including average total spending, purchase frequency, and order value.

**Explanation**

* The query evaluates customer risk by combining purchase history with demographic information.
* The customer\_profile section identifies each customer’s risk level by assessing purchase recency and total spending. In the risk\_summary section, customers are grouped by region and risk profile, providing insights on regional differences in customer engagement and spending habits.
* This approach highlights areas where customers may be more likely to churn or reduce spending.

**Code**

**A screenshot of a computer code

Description automatically generatedA table with numbers and letters

Description automatically generated**

**Result**

The output includes:

* Number of customers per region and risk profile.
* Average total spending, purchase frequency, and order value by risk level and location.
* This helps identify regions with higher churn risks, enabling targeted customer retention strategies.

1. Customer Lifetime Value Modeling: How can you leverage customer data (tenure, purchase history, engagement) to predict the lifetime value of different customer segments? This could inform targeted marketing and loyalty program strategies. Can you observe any common characteristics or purchase patterns among customers who have stopped purchasing?

**Approach**

* Customer Lifecycle Analysis: Calculate tenure, total spending, and classify customers by status (Active or Churn) and segment (Long-term or Short-term).
* Segment and Churn Analysis: Group data by customer segment, status, and region, computing lifetime value and predicted annual value.

**Explanation**

* The customer\_lifeStyle\_analysis step identifies customer tenure, spending behavior, and predicted annual value, categorizing each customer as "Churn" or "Active" based on recent activity.
* This helps identify customer segments (long- vs. short-term) and provides a lifetime value calculation.
* The segment\_analysis and churn\_analysis steps then summarize data by customer segment, engagement level, and location to analyze common patterns in high-value or at-risk customers.

**Code**

**A computer screen shot of a program code

Description automatically generatedA screenshot of a computer program

Description automatically generatedA screenshot of a computer

Description automatically generated**

**Result**

This output provides:

* Customer count, average tenure, lifetime value, and predicted annual value by segment and status.
* Counts and average lifetime value of churned customers by region and segment.
* These insights aid in identifying high-value customer profiles for targeted marketing and understanding characteristics linked to churn for retention strategies.

1. If data on promotional campaigns (discounts, events, email marketing) is available, how could you measure their impact on customer acquisition, retention, and overall sales?

To measure the impact of promotional campaigns on customer acquisition, retention, and overall sales, we can employ the following strategies:

* Customer Acquisition Analysis: Track the number of new customers gained during each campaign period compared to a baseline period without promotions. This provides insights into which campaigns are most effective in driving new customer interest.
* Retention and Repeat Purchases: Analyze the retention rate by monitoring how many customers make repeat purchases after a campaign ends. Compare this data to retention rates outside of campaign periods to assess whether promotions improve customer loyalty and long-term engagement.
* Sales Uplift Measurement: Compare total sales during campaign periods with average sales in similar non-campaign periods. This helps quantify the direct impact of the promotion on revenue growth, accounting for seasonal or other external influences where possible.
* Segmented Response Analysis: Break down the campaign impact by customer segments (e.g., by age, region, spending habits) to identify which groups respond best to specific types of promotions (like discounts vs. events). This enables more targeted and efficient campaign planning in the future.

Each of these measurements allows for a comprehensive view of how well promotional efforts drive desired customer behaviors and sales outcomes.

1. How would you approach this problem, if the objective and subjective questions weren't given?

* Data Quality Assessment: Begin by ensuring the data is clean, checking for any null values or duplicates. This step is crucial to maintain the integrity of the analysis.
* Key Performance Indicators (KPIs) Analysis: Identify relevant KPIs that align with the objectives of the analysis. This includes metrics related to customer purchases, churn rates, and overall sales performance.
* Question Generation: Generate insightful questions based on the available data, focusing on customer behavior, purchase patterns, and preferences.

Examples might include:

1. What are the most purchased albums or tracks?
2. Which genres have the highest sales?
3. How do purchasing behaviors vary by country?

* Customer Segmentation: Analyze customer data by segmenting it according to different criteria such as demographics, purchase history, and engagement levels. This allows for a more targeted approach to understanding customer behavior.
* Churn Analysis: Devote time to identifying factors contributing to customer churn. Examine which customer segments are more likely to stop purchasing and the reasons behind it.
* Comparative Analysis: Conduct comparative analyses between different countries to determine variations in customer preferences, spending habits, and the popularity of various music tracks, albums, or genres.
* Sales Insights: Identify and analyze the top-selling albums or tracks, focusing on factors such as pricing, promotion strategies, and customer demographics that may influence purchasing decisions.

1. How can you alter the "Albums" table to add a new column named "ReleaseYear" of type INTEGER to store the release year of each album?

**Approach**

* Alter Table: Execute an ALTER TABLE command to add a new column named ReleaseYear of type INTEGER to the Albums table.
* Update Records: Use multiple UPDATE statements to assign the release year for specific albums based on their album\_id.

**Explanation**

* The ALTER TABLE statement modifies the structure of the Albums table by adding the ReleaseYear column, allowing storage of each album's release year.
* The subsequent UPDATE statements set the ReleaseYear to either 2017 or 2018 for specific albums identified by their album\_id.

**Code**

**A screenshot of a computer program

Description automatically generatedA screenshot of a computer program

Description automatically generatedA screenshot of a music list

Description automatically generated**

**Result**

* After executing the SQL commands, the Albums table includes the new ReleaseYear column, with the specified albums updated to reflect their respective release years. The table can now accurately store and display the release year for each album.

1. Chinook is interested in understanding the purchasing behavior of customers based on their geographical location. They want to know the average total amount spent by customers from each country, along with the number of customers and the average number of tracks purchased per customer. Write an SQL query to provide this information.

**Approach**

* Calculate Tracks per Customer: Create a Common Table Expression (CTE) called tracks\_per\_customer to sum the quantity of tracks purchased by each customer.
* Calculate Customer Spending: Create another CTE named customer\_spending that sums the total spent by customers grouped by their country, combining data from the customer, invoice, and the previous CTE.

**Explanation**

* The tracks\_per\_customer CTE calculates the total number of tracks purchased by each customer by summing the quantity from the invoice\_line table.
* The customer\_spending CTE aggregates total spending for each customer and associates it with their respective country.

**Code**

**A screenshot of a computer program

Description automatically generatedA table with numbers and numbers

Description automatically generated**

**Result**

* The resulting dataset provides insights into the purchasing behavior of customers, highlighting which countries have the highest average spending and track purchases per customer.