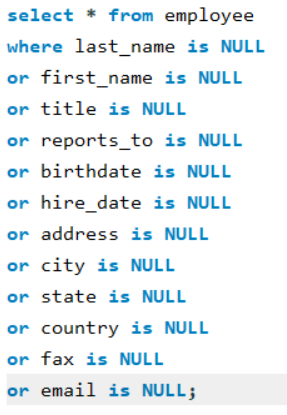
**Learners have to come up with a Report to support the answers to the following questions and suggestions**

Objective Questions

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

Ans :

For handling missing values I used the following SQL query to identify rows where any of the columns might contain a null value:

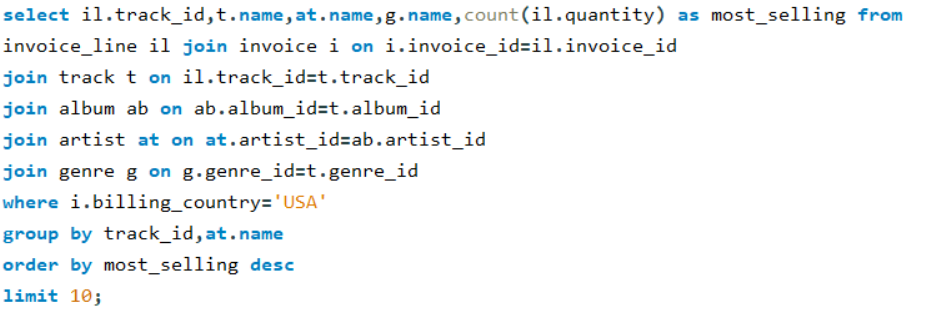


The data provided possess 0 duplicates although there are missing values in 3 tables.

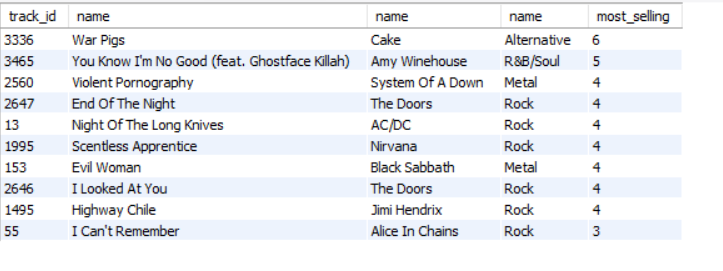
By executing this query the reports\_to column contained null values which I handled by replacing the null values with NA using coalesce function

Similary I applied the same approach to other table also to handle Null Values

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

Ans: 

**OUTPUT :**



**Insights:**

Rock is the most preferred genre in USA.

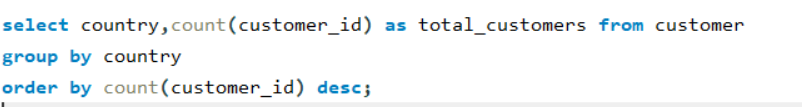
**Top Selling Tracks:** "War Pigs" by Cake and "You Know I'm No Good (feat. Ghostface Killah)" by Amy Winehouse are the top-selling tracks, each selling 6 and 5 units respectively.

**Recommendations:**

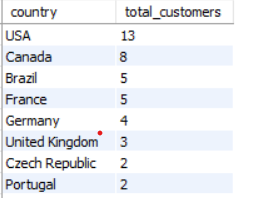
* **Focus on High-Demand Genres:** Allocate more resources to promote and stock Rock music, ensuring availability aligns with customer preferences.
* **Targeted Marketing Campaigns:** Develop genre-focused advertising strategies to engage specific audience segments, such as Rock enthusiasts or Jazz aficionados.
* **Explore Genre Expansion:** Gradually introduce underrepresented genres to test their market potential while maintaining a strong presence in popular categories.

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

Ans :



**OUTPUT :**



**Insights:**

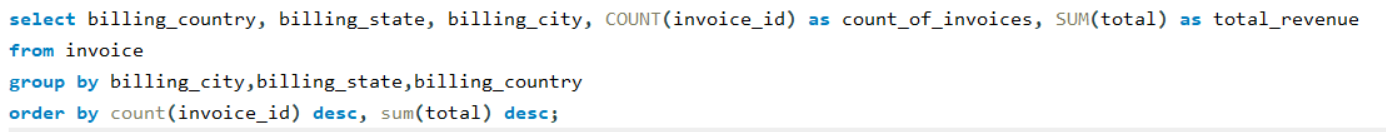
* Customers are distributed across 24 different countries, showcasing a global presence.
* The United States accounts for the largest proportion of customers.

**Recommendations:**

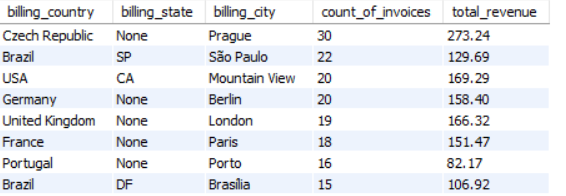
* Prioritize efforts in the U.S. market to leverage its significant customer base.
* Identify and capitalize on expansion opportunities in high-potential markets such as Canada, Brazil, France, and Germany.
* Develop region-specific marketing strategies to cater to the preferences and cultural nuances of different countries.

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

Ans :



**OUTPUT:**



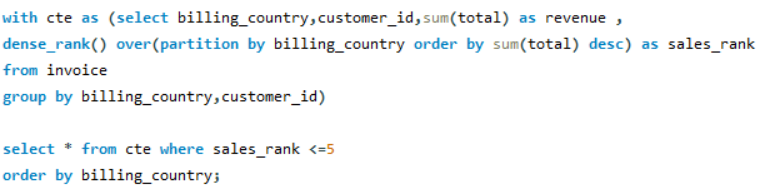
**Insights:**

* Prague, Czech Republic, leads in revenue generation among all locations.
* São Paulo, Brazil, has a large number of invoices but comparatively lower revenue.

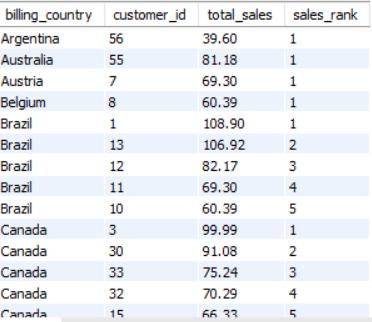
**Recommendations:**

* Conduct an in-depth analysis of Prague's revenue performance to uncover factors contributing to its success and apply similar strategies in other cities.
* Assess revenue per invoice across all regions to pinpoint inefficiencies and refine pricing or service strategies to maximize profitability.

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



**OUTPUT :**

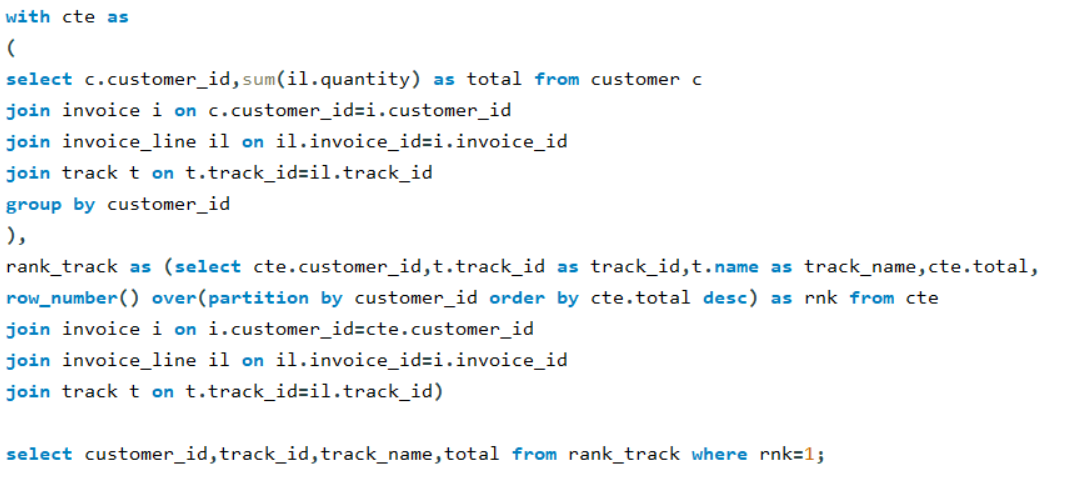


* Leading customers differ across countries, highlighting regional variations in customer behavior.

**Recommendations:**

* **Targeted Outreach:** Develop customized marketing initiatives for high-value customers in each country to enhance engagement.
* **Incentive-Based Programs:** Introduce loyalty schemes with tiers linked to spending levels to encourage repeat business.
* **Regional Customer Segmentation:** Categorize customers by spending habits and geographical location to enable tailored communication strategies.
* **Behavioral Analysis:** Examine purchasing patterns and preferences of top customers to identify key factors driving their high-value transactions.

1. Identify the top-selling track for each customer



**OUTPUT :**

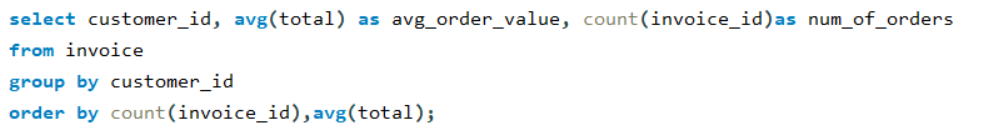


**Insights:**

1. **Customer-Specific Preferences**:
   * Each customer has a unique top-selling track, highlighting their diverse preferences.
   * For example, Customer 3003's top track is *"All I Want Is You"*, while Customer 3424 prefers *"Turandot, Act III, Nessun dorma!"*.
2. **Popularity Distribution**:
   * The total purchases for the top tracks vary significantly, ranging from as high as 146 (*"Turandot, Act III, Nessun dorma!"*) to as low as 38 (*"Mr. Moon"*).
   * This indicates a disparity in customer demand for certain tracks.
3. **Genre Influence**:
   * Tracks such as *"Welcome to the Jungle"* and *"Cochise"* may appeal to specific genres (like rock), suggesting a genre-specific popularity.
4. **High-Value Customers**:
   * Customers like 3424 (*146 total*) and 1479 (*130 total*) stand out as high-value customers based on their significant purchases of top tracks.

**Recommendations:**

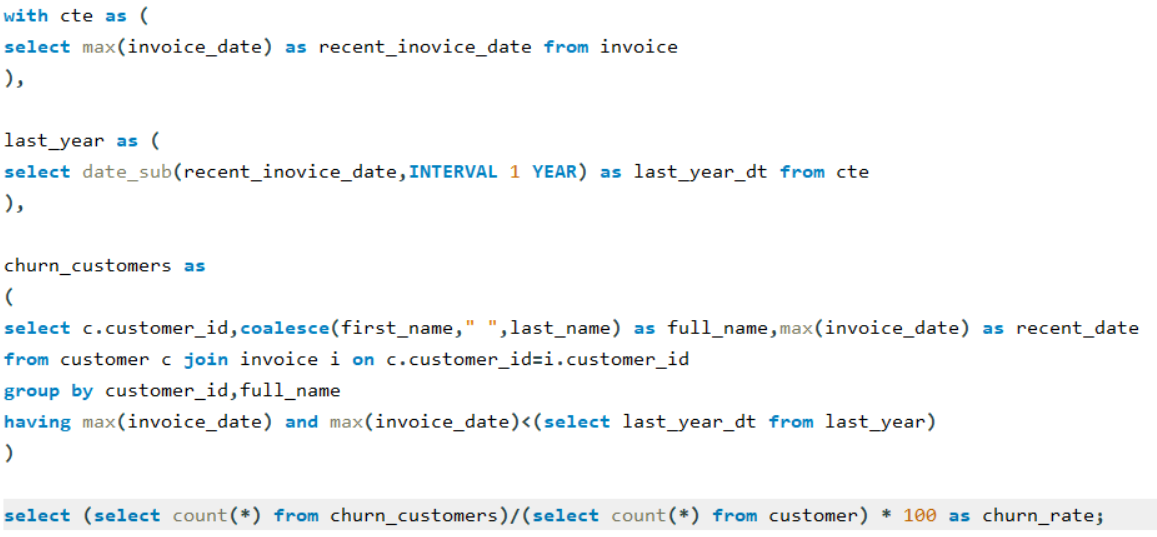
1. **Personalized Offers**:
   * Use customer-specific purchase data to create personalized playlists or recommend similar tracks to maximize sales.
2. Are there any patterns or trends in customer purchasing behavior (e.g., frequency of purchases, preferred payment methods, average order value)?



**Insights:**

**Focus on High-Value and High-Frequency Customers**: Prioritize customers with high average order values and frequent purchases by offering exclusive loyalty rewards, personalized offers, or early access to premium products. This will ensure sustained revenue from your top-performing customers.

1. What is the customer churn rate?



**OUTPUT :**

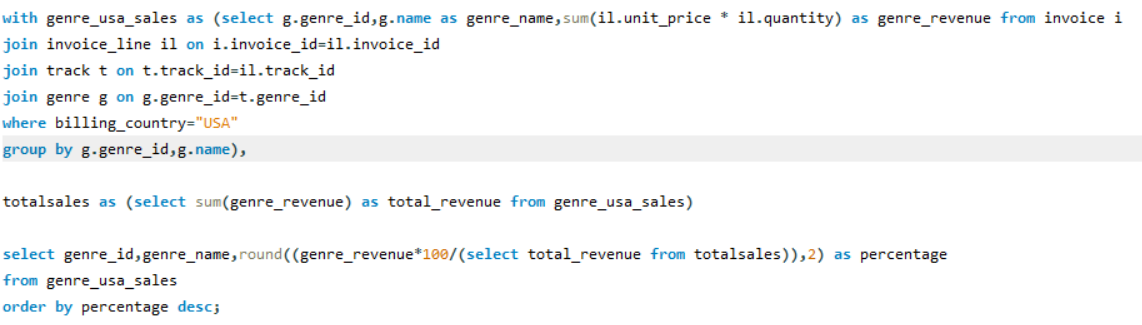
# churn\_rate

'1.6949'

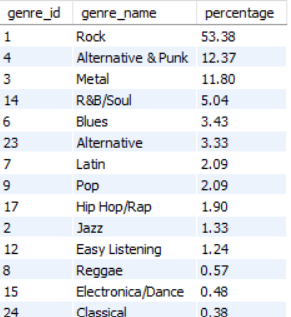
**Insights:**

* The customer churn rate stands at 1.6949, meaning nearly 1.7% of customers leave during a given time frame.
* A churn rate below 2% reflects strong customer retention and minimal turnover.
* Sustaining such a low churn rate is vital for fostering long-term growth and building customer loyalty.
* To reduce churn even further, businesses can focus on boosting customer engagement, elevating service quality, and actively addressing customer concerns and feedback.

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



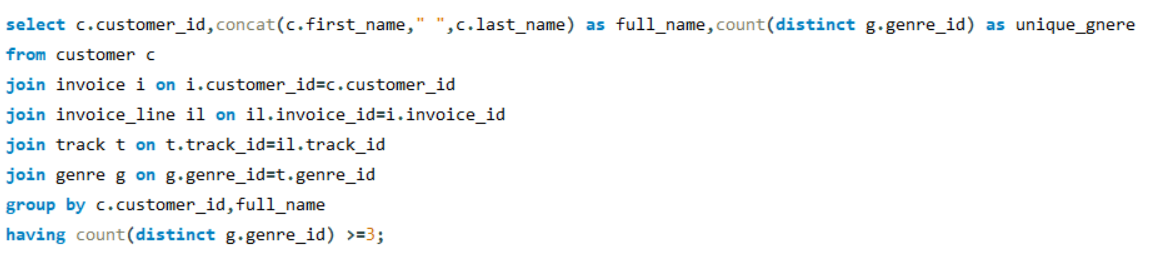
**OUTPUT:**



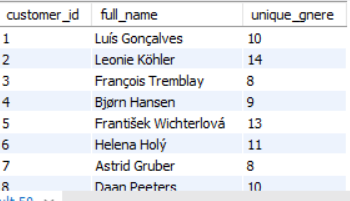
**Insights**

* **Rock Leads the Market**: Rock music contributes to over half of all sales in the U.S., making it the most popular genre.
* **Cake’s Popularity**: The artist Cake has a significant fan base, with their music being a top choice for many customers, earning them the title of best-selling artist.

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



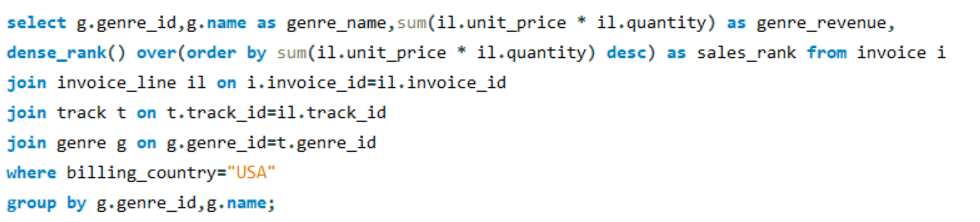
**OUTPUT :**

****

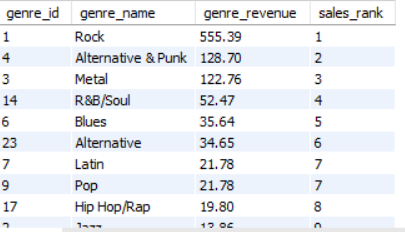
**Insights**

* Varied Preferences: Every customer has explored and purchased music from a minimum of three different genres, showcasing diverse tastes.
* Most Versatile Customer: Leonie Köhler stands out for buying music from the highest number of genres, totaling 14.

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



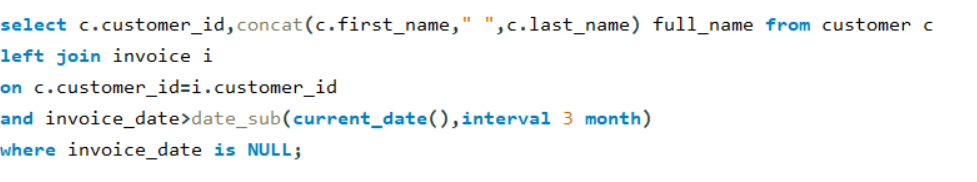
**OUTPUT:**

****

**Insights**

* Rock Dominates: Rock is the most popular genre, significantly outperforming others.
* Preference for Guitar-Centric Styles: The top three genres—Rock, Alternative & Punk, and Metal—highlight a strong customer inclination toward guitar-driven music.

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

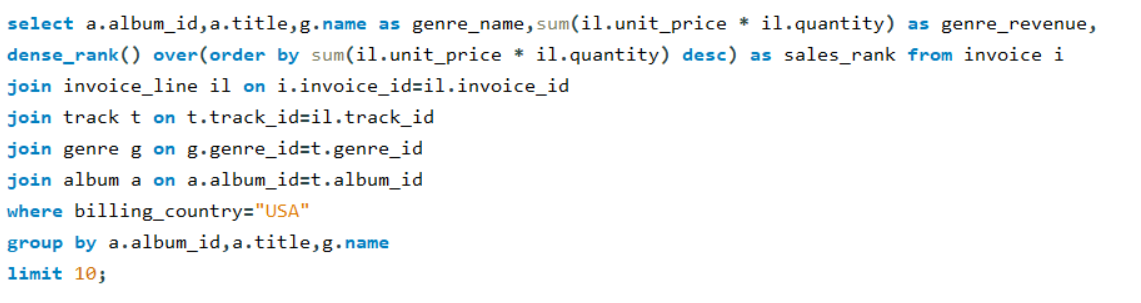


**OUTPUT:**

****

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.



**OUTPUT:**



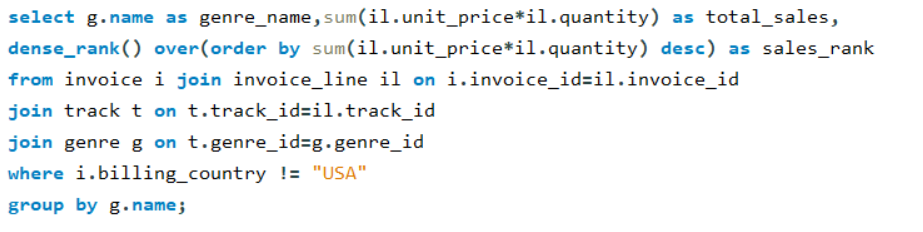
**Insights**

* **Rock Dominates the Market**: Rock remains the most popular genre among US customers.
* **Top Rock Albums**: The leading albums in the Rock category are *"From The Muddy Banks Of The Wishkah [live]"*, *"Are You Experienced?"*, and *"The Best Of Van Halen, Vol. I"*.

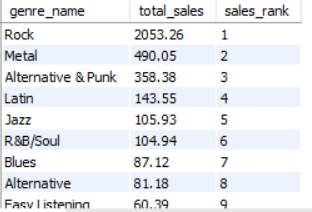
**Recommendations**

* **Focus on Promotions**: Concentrate marketing campaigns on these three top-selling Rock albums to drive sales in the US market.
* **Introduce Bundling Offers**: Create exclusive bundles or promotional deals that feature these albums to attract more customers.
* **Leverage Targeted Ads**: Use purchase data to pinpoint US customers with an interest in Rock music and promote these albums specifically to them.

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



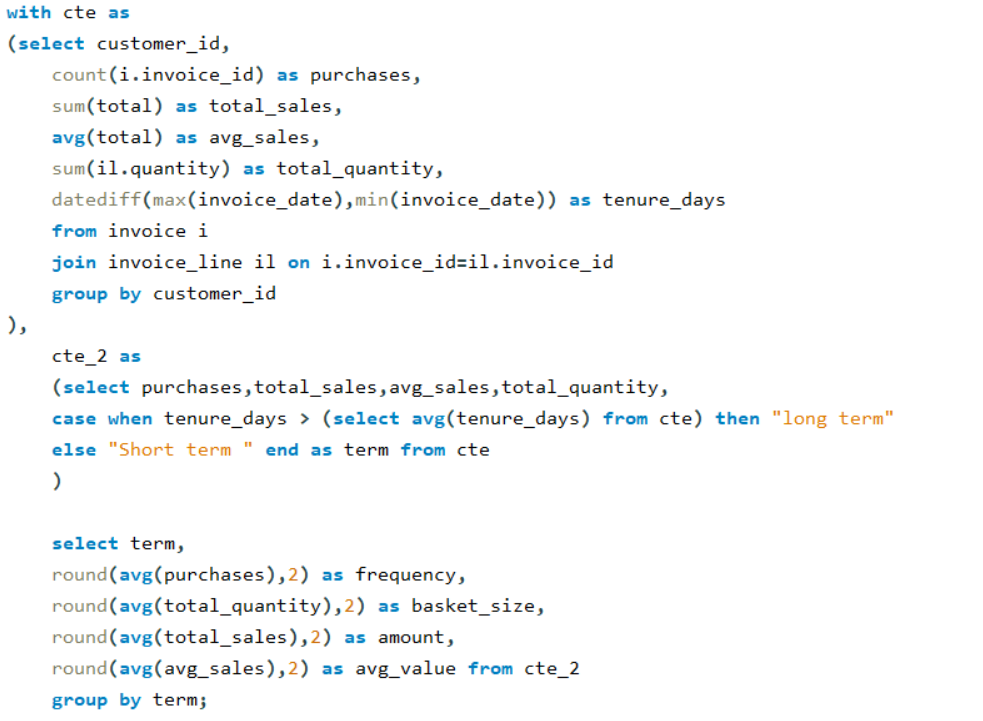
**OUTPUT:**



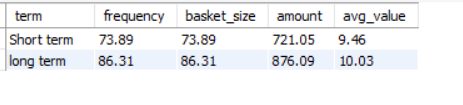
**Insights**

* **Global Dominance of Rock**: Rock remains a favorite genre both in the USA and globally.
* **Regional Preferences**: While Rock is widely popular, the US market exhibits distinct genre preferences compared to international markets.
* **Growing Genres Internationally**: Metal, Latin, and Jazz are gaining significant traction outside the USA, showing strong performance in emerging markets.

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?

Ans : 

**OUTPUT :**



**Insights :**

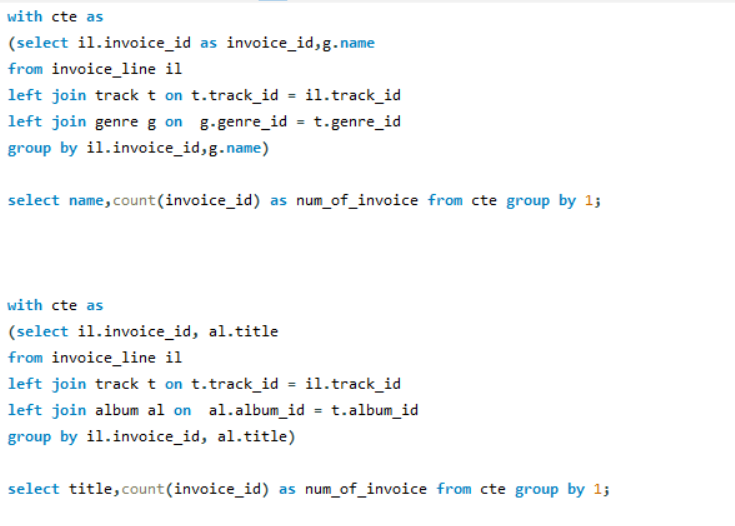
Revenue generated by long-term customers surpasses that of short-term customers, highlighting their importance for sustained business growth.

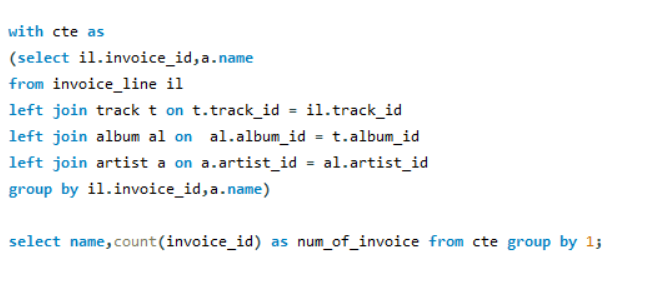
**Recommendations**

* **Strengthen Retention Strategies**: Develop loyalty programs or personalized offers to encourage customers to remain engaged for the long term.
* **Focus on Value Enhancement**: Provide exclusive benefits or incentives for long-term customers to further boost their average order value and loyalty.

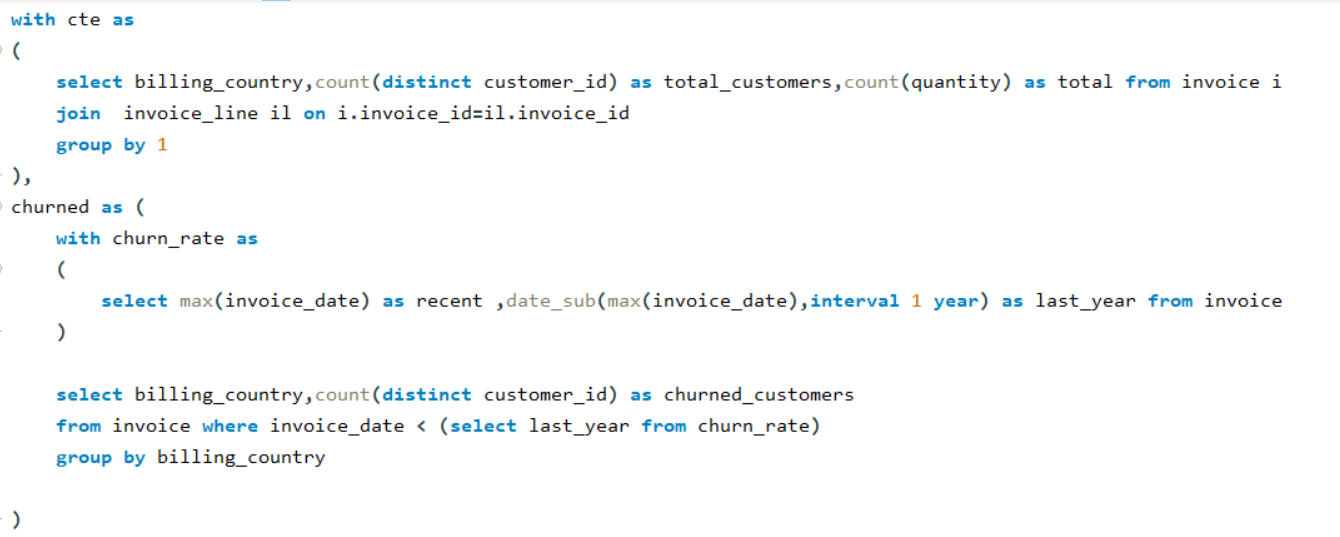
1. Product Affinity Analysis: Which music genres, artists, or albums are frequently purchased together by customers? How can this information guide product recommendations and cross-selling initiatives?

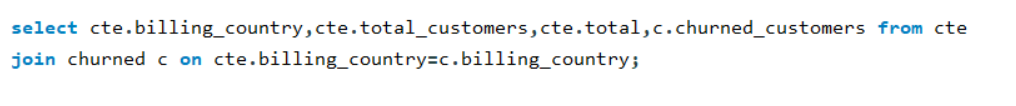
Ans :

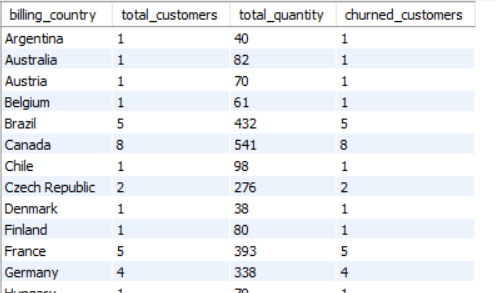




1. Regional Market Analysis: Do customer purchasing behaviors and churn rates vary across different geographic regions or store locations? How might these correlate with local demographic or economic factors?







**Customer Purchasing Behaviors**

1. **High Spending and Frequency**:
   * Prague ($9.38, 15), Mountain View ($8.60, 10), London ($8.72, 9.5).  
     *Focus on premium products and personalized marketing.*
2. **Moderate Spending and Frequency**:
   * São Paulo ($5.98, 11), Berlin ($7.94, 10).  
     *Boost customer loyalty and repeat purchases.*
3. **Low Spending and Frequency**:
   * Edmonton ($2.97, 10), Copenhagen ($3.76, 10).  
     *Introduce promotions and assess local factors.*

**Churn Analysis**

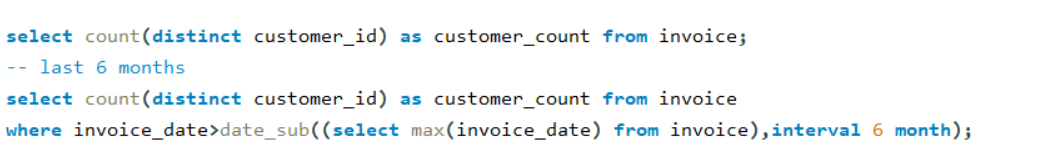
* **High Churn Regions**: São José dos Campos, Montréal, São Paulo, and others.  
  *Implement loyalty programs and gather customer feedback to improve retention.*

**Recommendations**

* **Affluent Areas**: Offer exclusive, high-value products.
* **Emerging Markets**: Focus on customer engagement and value-driven offers.
* **High Churn Areas**: Provide discounts, better service, and retention incentives.

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?

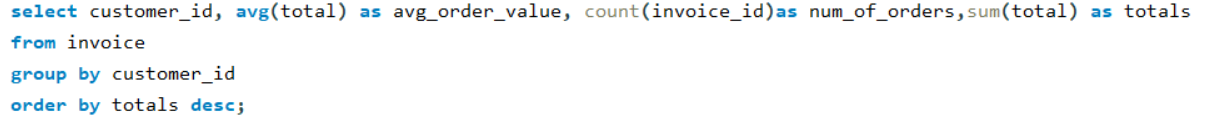
Ans :

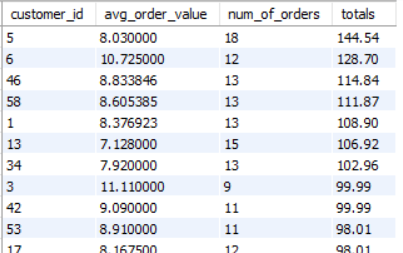


From 59 total\_customers

There were only 43 customers in last 6 months

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?





**Recommedations:**

Customer data, including purchase history, tenure, and engagement, can be leveraged to identify traits like loyalty, high-value or low-value status, and buying frequency.

Recent high-value customers show potential for long-term loyalty and should be prioritized for targeted marketing and promotional campaigns to nurture this potential.

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?

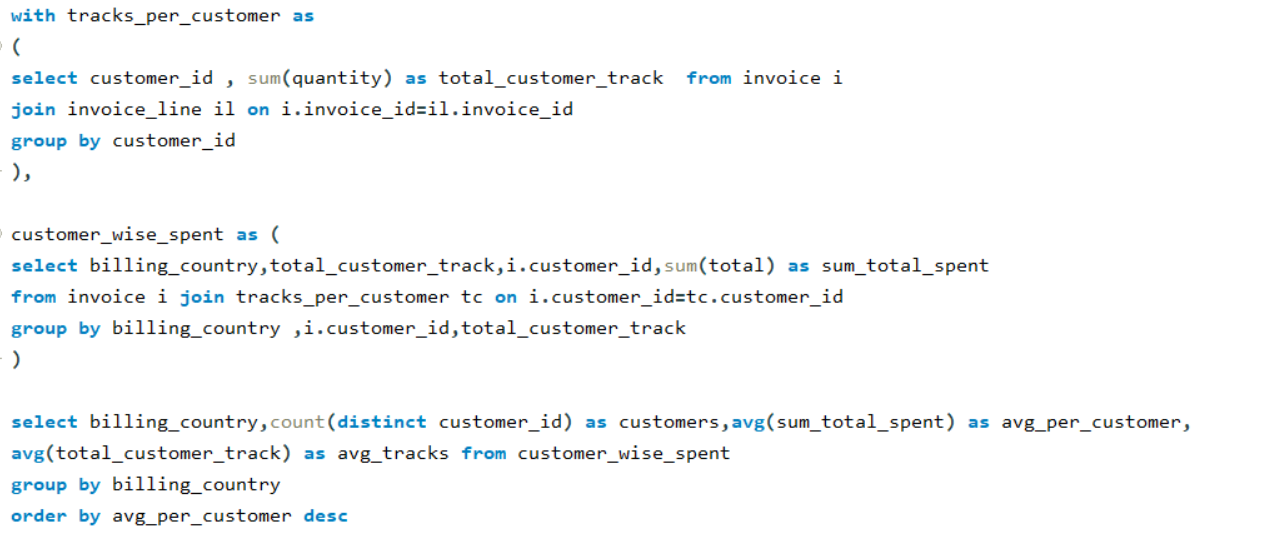
Ans :

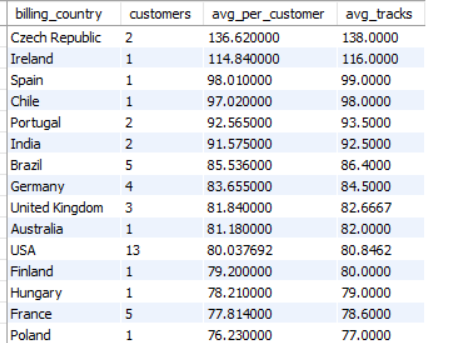
* **Customer Acquisition**: Count new customers acquired during and after the campaign period.
  + Metric: New customers (e.g., count of unique customer IDs).
* **Customer Retention**: Measure the number of existing customers who made repeat purchases.
  + Metric: Repeat purchases, retention rate (e.g., percentage of returning customers).
* **Overall Sales Impact**: Assess total revenue, order frequency, and average order value.
  + Metric: Total sales, frequency, and average order value before, during, and after the campaign.

1. How would you approach this problem, if the objective and subjective questions weren't given?
2. **Understand the Problem Context**
   * Gather comprehensive details about the dataset, including its structure, attributes, and relevance to business objectives.
   * Clearly define the goals of the analysis and the key questions it aims to address (e.g., understanding customer behavior, identifying trends, etc.).
3. **Data Exploration**
   * Conduct an exploratory analysis to uncover patterns, trends, and outliers in the data.
   * Focus on key metrics such as customer purchase frequency, average order value, tenure, and churn rates to gain an overview of customer behavior.
4. **Customer Segmentation**
   * Group customers based on distinct characteristics like spending patterns, frequency of purchases, and level of engagement.
   * Create categories such as high-value customers, loyal buyers, and occasional shoppers to enable targeted analysis.
5. **Derive Insights**
   * Investigate trends and relationships to extract actionable insights, such as identifying customer groups that drive the majority of revenue.
   * Pinpoint factors affecting customer retention, loyalty, or reduced activity to guide decision-making.
6. 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?

Ans : alter table album add column Releaseyear int;

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.





**Insights:**

1. **High-Spending Regions**: Customers in **Czech Republic** ($136.62) and **Ireland** ($114.84) spend the most and buy the most tracks.
2. **Large Volume Market**: The **USA** has the most customers (13) but relatively low average spending ($80.04).
3. **Emerging Markets**: **Brazil** and **Germany** show moderate spending with room for growth.
4. **Low-Spending Regions**: **Poland** and **France** have the lowest spending per customer.

**Recommendations:**

1. **Prioritize High-Spenders**: Focus premium offers and promotions in **Czech Republic** and **Ireland**.
2. **Boost USA Engagement**: Target the USA with campaigns to increase spending and track purchases.
3. **Develop Emerging Markets**: Invest in customer retention strategies for **Brazil** and **Germany**.
4. **Improve Low-Spend Regions**: Offer localized discounts or bundles in **Poland** and **France**.