**Chinook Music Store Analysis**

Objective Questions

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

SQL Query:

SELECT \* FROM album;

SELECT \* FROM artist;

SELECT COUNT(\*) FROM customer -- 49 company, 29 state, 47 fax values are null in the customer table

WHERE fax is NULL;

SELECT \* from employee; -- 1 reports\_to value is null in the employee table

SELECT \* FROM genre;

SELECT \* FROM invoice;

SELECT \* FROM invoice\_line;

SELECT \* FROM media\_type;

SELECT \* FROM playlist;

SELECT \* FROM playlist\_track;

SELECT COUNT(\*) FROM track -- 978 composer columns are null in the track table

WHERE composer is NULL

Yes, there were a number of missing values in the dataset like: -

* 49 company, 29 state, 47 fax values are null in the customer table.
* 1 reports\_to value is null in the employee table.
* 978 composer values are null in the track table.

There are no duplicate values in the whole dataset. In case of null values I would use COALESCE function to handle the situation

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

SQL Query: -

SELECT Top\_selling\_track, Top\_artist, Top\_genre FROM

(

SELECT t.name Top\_selling\_track, a.name Top\_artist, g.name Top\_genre, SUM(t.unit\_price \* il.quantity) FROM track t

LEFT JOIN invoice\_line il on t.track\_id = il.track\_id

LEFT JOIN invoice i on i.invoice\_id = il.invoice\_id

LEFT JOIN album al on al.album\_id = t.album\_id

LEFT JOIN artist a on a.artist\_id = al.artist\_id

LEFT JOIN genre g on g.genre\_id = t.genre\_id

WHERE billing\_country = "USA"

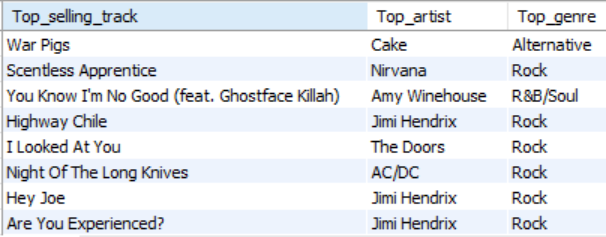
GROUP BY t.name, a.name, g.name

ORDER BY SUM(total) DESC

LIMIT 10

) Agg\_table;

OUTPUT: -



Therefore, the top genre in USA is Rock.

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

SQL Query: -

SELECT city, country, COUNT(customer\_id) FROM customer

GROUP BY 1,2

ORDER BY country;

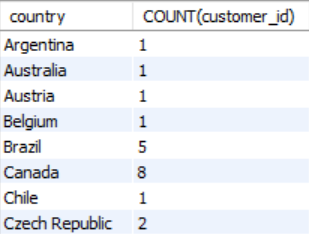
SELECT country, COUNT(customer\_id) FROM customer

GROUP BY 1

ORDER BY 1;

SELECT COUNT(distinct country) FROM customer;

OUTPUT: -

* The Customer Demographic Breakdown based on location is very diversified.
* The Chinook's customer base is spread over 24 countries with max number of customers from USA.
* The customer table does not have age and gender columns to understand the customer breakdown

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

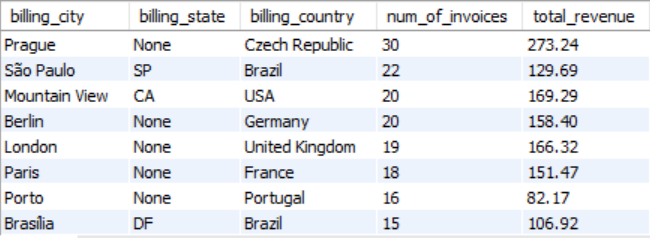
SQL Query: -

SELECT billing\_city, billing\_state, billing\_country, COUNT(invoice\_id) num\_of\_invoices, SUM(total) total\_revenue FROM invoice

GROUP BY 1,2,3

ORDER BY COUNT(invoice\_id) DESC, SUM(total) DESC

OUTPUT: -



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

SQL Query: -

WITH cte as

(

SELECT country, first\_name, last\_name, SUM(t.unit\_price \* il.quantity) total\_revenue FROM customer c

LEFT JOIN invoice i on i.customer\_id = c.customer\_id

LEFT JOIN invoice\_line il on il.invoice\_id = i.invoice\_id

LEFT JOIN track t on t.track\_id = il.track\_id

GROUP BY 1,2,3

ORDER BY country

),

cte2 as

(

SELECT country, first\_name, last\_name,

RANK() OVER(PARTITION BY country ORDER BY total\_revenue DESC) rk

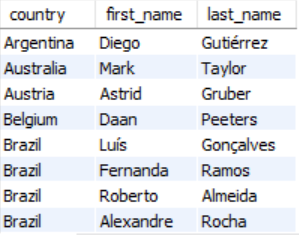
FROM cte

)

SELECT country, first\_name, last\_name FROM cte2

WHERE rk <= 5;

OUTPUT: -



1. **Identify the top-selling track for each customer**

SQL Query: -

SELECT first\_name, last\_name, t.name Track\_name, SUM(quantity) Total\_quantity FROM customer c

LEFT JOIN invoice i on i.customer\_id = c.customer\_id

LEFT JOIN invoice\_line il on il.invoice\_id = i.invoice\_id

LEFT JOIN track t on t.track\_id = il.track\_id

GROUP BY 1,2,3

ORDER BY SUM(quantity) DESC;

OUTPUT: -



\*The customers had similar quantity of different tracks. So its was difficult to select one as the top track for each customer.

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

SQL Query :

SELECT customer\_id, COUNT(invoice\_id) num\_invoices, AVG(total) avg\_sales FROM invoice

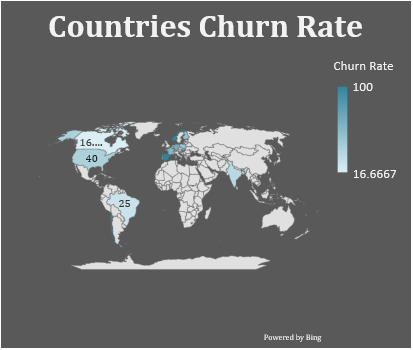
GROUP BY 1

ORDER BY COUNT(invoice\_id) DESC, AVG(total) DESC

OUTPUT :



No there is no correlation or trend between the number/frequency of orders by different customers and the average sales generated by these customers. The average sales most probably depend on the unit price of each track and not the number of orders.



1. **What is the customer churn rate?**

Customer churn rate is defined as the percentage of customers who stop doing business with an organization over a period of time.

SQL Query: -

WITH num\_cust\_in\_1st\_3months as

(

SELECT COUNT(customer\_id) ttl from invoice

WHERE invoice\_date BETWEEN '2017-01-01' AND '2017-03-31'

),-- I have taken the assumption that total number of customers in the beginning is equal to the customers joining in the first 3 months.

num\_cust\_in\_last\_2months as

(

SELECT COUNT(customer\_id) l\_num FROM invoice

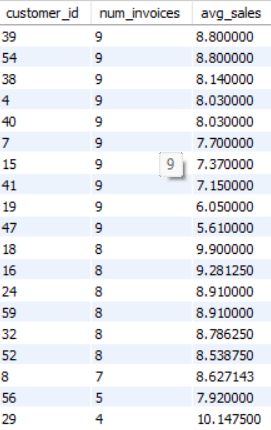
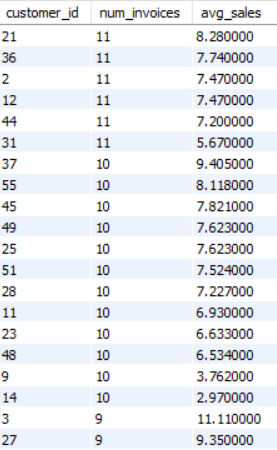
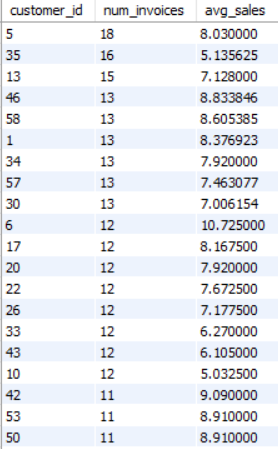
WHERE invoice\_date BETWEEN '2020-11-01' AND '2020-12-31'

) -- I have taken the assumption that churn rate will be calculated on the basis of the number of customers left in the last two months.

SELECT ((SELECT ttl FROM num\_cust\_in\_1st\_3months)-(SELECT l\_num FROM num\_cust\_in\_last\_2months))/(SELECT ttl FROM num\_cust\_in\_1st\_3months) \* 100 as churn\_rate;

OUTPUT: -





Therefore, the customer churn rate of the company is 40.8163 based on the total number of customers in first 3 months i.e 49 and the number of customers present in the last 2 months i.e 29. So, number of customers lost = 49-29 = 20.

Churn rate = 20/49 = 40.8163

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

SQL Query: -

WITH cte as

(

SELECT SUM(total) total\_revenue\_for\_USA FROM invoice

WHERE billing\_country = 'USA'

),

genre\_sales as

(

SELECT g.genre\_id, g.name, sum(t.unit\_price \* il.quantity) total\_revenue\_for\_genre FROM track t

LEFT JOIN genre g on g.genre\_id = t.genre\_id

LEFT JOIN invoice\_line il on il.track\_id = t.track\_id

LEFT JOIN invoice i on i.invoice\_id = il.invoice\_id

WHERE billing\_country = 'USA'

GROUP BY 1,2

ORDER BY total\_revenue\_for\_genre DESC

),

ranking as

(

SELECT genre\_id, name, ROUND(total\_revenue\_for\_genre/(SELECT total\_revenue\_for\_USA FROM cte) \* 100,2) percentage\_contribution,

DENSE\_RANK() OVER(ORDER BY ROUND(total\_revenue\_for\_genre/(SELECT total\_revenue\_for\_USA FROM cte) \* 100,2) DESC) rk FROM genre\_sales

)

SELECT ranking.genre\_id, ranking.name genre\_name, a.name artist\_name, percentage\_contribution, rk FROM ranking

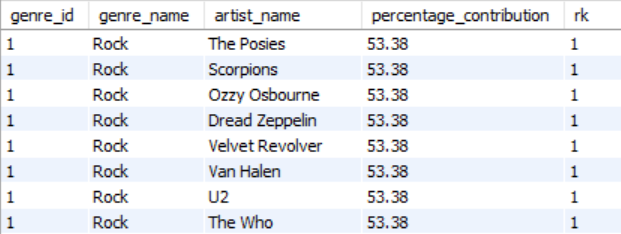
LEFT JOIN track t on t.genre\_id = ranking.genre\_id

LEFT JOIN album al on al.album\_id = t.album\_id

LEFT JOIN artist a on a.artist\_id = al.artist\_id

GROUP BY 1,2,3,4

OUTPUT: -



Therefore the top selling genre in USA is Rock.

And

The Posies, Scorpions, Ozzy Osbourne, Dread Zeppelin, Velvet Revolver, Van Halen, U2, The Who, The Rolling Stones, The Police, The Doors, The Cult, Terry Bozzio, Tony Levin & Steve Stevens, Stone Temple Pilots, Soundgarden, Skank, Lenny Kravitz, Santana, Rush, Red Hot Chili Peppers, Raul Seixas, R.E.M., Queen, Pink Floyd, Pearl Jam, Paul D'Ianno, Page & Plant, O Terço, Nirvana, Men At Work, Marillion, Led Zeppelin, Kiss, Joe Satriani, Jimi Hendrix, Jamiroquai, Iron Maiden, Guns N' Roses, Foo Fighters, Faith No More, Def Leppard, Deep Purple, Creedence Clearwater Revival, David Coverdale, Frank Zappa & Captain Beefheart, Audioslave, Alice In Chains, Alanis Morissette, Aerosmith, AC/DC, Accept

are all the top artists who are associated with the Rock genre.

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

SQL Query: -

SELECT name\_of\_customer FROM

(

SELECT CONCAT(first\_name, ' ', last\_name) name\_of\_customer, COUNT(DISTINCT g.name) FROM customer c

LEFT JOIN invoice i on i.customer\_id = c.customer\_id

LEFT JOIN invoice\_line il on il.invoice\_id = i.invoice\_id

LEFT JOIN track t on t.track\_id = il.track\_id

LEFT JOIN genre g on g.genre\_id = t.genre\_id

GROUP BY 1 HAVING COUNT(DISTINCT g.name) >= 3

ORDER BY COUNT(DISTINCT g.name) DESC

) agg\_table

OUTPUT: -



Leonie Köhler is the person who has bought tracks from 14 different genres.

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

SQL Query: -

WITH cte as

(

SELECT t.genre\_id, g.name, SUM(t.unit\_price \* il.quantity) sale\_performance FROM track t

LEFT JOIN genre g on g.genre\_id = t.genre\_id

LEFT JOIN invoice\_line il on il.track\_id = t.track\_id

LEFT JOIN invoice i on i.invoice\_id = il.invoice\_id

WHERE billing\_country = 'USA'

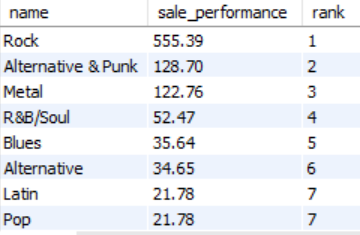
GROUP BY 1, 2

)

SELECT name, sale\_performance,

DENSE\_RANK() OVER(ORDER BY sale\_performance DESC) `rank` FROM cte;

OUTPUT: -



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

SQL Query: -

WITH last\_3\_months as

(

SELECT \* from invoice

WHERE invoice\_date > (SELECT MAX(invoice\_date) FROM invoice) - INTERVAL 3 MONTH

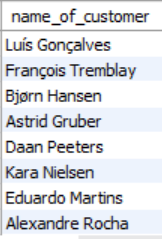
)

SELECT CONCAT(first\_name, ' ', last\_name) name\_of\_customer FROM customer c

LEFT JOIN last\_3\_months lm on lm.customer\_id = c.customer\_id

WHERE invoice\_id is NULL;

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.**

SQL Query: -

WITH genre\_sales as

(

SELECT g.genre\_id, g.name, sum(t.unit\_price \* il.quantity) total\_revenue\_for\_genre FROM track t

LEFT JOIN genre g on g.genre\_id = t.genre\_id

LEFT JOIN invoice\_line il on il.track\_id = t.track\_id

LEFT JOIN invoice i on i.invoice\_id = il.invoice\_id

WHERE billing\_country = 'USA'

GROUP BY 1,2

ORDER BY total\_revenue\_for\_genre DESC

),

ranking as

(

SELECT genre\_id, name, total\_revenue\_for\_genre,

DENSE\_RANK() OVER(ORDER BY total\_revenue\_for\_genre DESC) rk FROM genre\_sales

),

genre\_album as

(

SELECT ranking.genre\_id, ranking.name genre\_name, al.title album\_name FROM ranking

LEFT JOIN track t on t.genre\_id = ranking.genre\_id

LEFT JOIN album al on al.album\_id = t.album\_id

LEFT JOIN artist a on a.artist\_id = al.artist\_id

WHERE rk = 1

GROUP BY 1,2,3

),

best\_album as

(

SELECT al.album\_id, title, SUM(t.unit\_price \* il.quantity) FROM album al

LEFT JOIN track t on t.album\_id = al.album\_id

LEFT JOIN invoice\_line il on il.track\_id = t.track\_id

GROUP BY 1,2

ORDER BY SUM(t.unit\_price \* il.quantity) desc

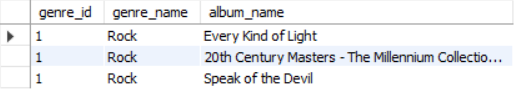
)

SELECT genre\_id, genre\_name, album\_name FROM genre\_album

inner join best\_album on best\_album.title = genre\_album.album\_name

LIMIT 3

OUTPUT:



Recommendation: - Top 3 albums that should be prioritised for advertisements and promotions in the USA based on genre analysis are

* Every Kind of Light
* 20th Century Masters - The Millennium Collection: The Best of Scorpions
* Speak of the Devil

All these albums are from the genre ROCK as it is the most popular genre in the USA.

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

SQL Query:

SELECT g.genre\_id, g.name, sum(t.unit\_price \* il.quantity) total\_revenue\_for\_genre FROM track t

LEFT JOIN genre g on g.genre\_id = t.genre\_id

LEFT JOIN invoice\_line il on il.track\_id = t.track\_id

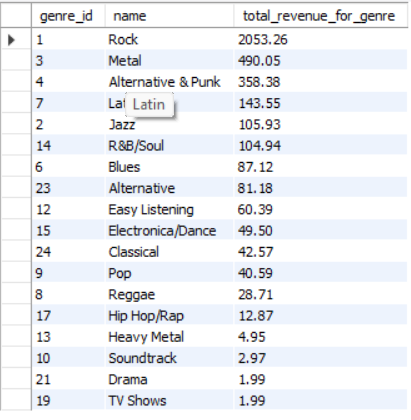
LEFT JOIN invoice i on i.invoice\_id = il.invoice\_id

WHERE billing\_country != 'USA'

GROUP BY 1,2

ORDER BY total\_revenue\_for\_genre DESC

OUTPUT :



The commonality between the data regarding USA and rest of the countries is that the ROCK genre has been taking the top spot in both the data. The 2nd and 3rd spot is taken by METAL & ALTERNATIVE & PUNK genre in the world respectively.

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?**

SQL Query: -

WITH cte as

(

SELECT i.customer\_id, MAX(invoice\_date), MIN(invoice\_date), abs(TIMESTAMPDIFF(MONTH, MAX(invoice\_date), MIN(invoice\_date))) time\_for\_each\_customer, SUM(total) sales, SUM(quantity) items, COUNT(invoice\_date) frequency FROM invoice i

LEFT JOIN customer c on c.customer\_id = i.customer\_id

LEFT JOIN invoice\_line il on il.invoice\_id = i.invoice\_id

GROUP BY 1

ORDER BY time\_for\_each\_customer DESC

),

average\_time as

(

SELECT AVG(time\_for\_each\_customer) average FROM cte

),-- 1244.3220 Days OR 40.36 Months

categorization as

(

SELECT \*,

CASE

WHEN time\_for\_each\_customer > (SELECT average from average\_time) THEN "Long-term Customer" ELSE "Short-term Customer"

END category

FROM cte

)

SELECT category, SUM(sales) total\_spending, SUM(items) basket\_size, COUNT(frequency) frequency FROM categorization

GROUP BY 1

OUTPUT :

A screenshot of a computer screen

Description automatically generated

Insights: - It can be seen that the spending amount and basket size for long-term customers is more than the short-term customers.

Recommendations: - It shows that customer loyalty plays an important role in increasing the revenue of the company because the long-term customers tend to buy more than the short-term customers. Therefore, the company should focus on the retention rate of the customers so as to increase the sales over time.

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?

SQL Query: -

WITH cte as

(

SELECT invoice\_id, COUNT(DISTINCT g.name) num FROM invoice\_line il

left JOIN track t on t.track\_id = il.track\_id

left JOIN genre g on g.genre\_id = t.genre\_id

GROUP BY 1 HAVING COUNT(DISTINCT g.name) > 1

)

SELECT cte.invoice\_id, num, g.name FROM cte

left join invoice\_line il on il.invoice\_id = cte.invoice\_id

left JOIN track t on t.track\_id = il.track\_id

left JOIN genre g on g.genre\_id = t.genre\_id

GROUP BY 1,2,3;

WITH cte as

(

SELECT invoice\_id, COUNT(DISTINCT al.title) num FROM invoice\_line il

left JOIN track t on t.track\_id = il.track\_id

left JOIN album al on al.album\_id = t.album\_id

GROUP BY 1 HAVING COUNT(DISTINCT al.title) > 1

)

SELECT cte.invoice\_id, num, al.title FROM cte

left join invoice\_line il on il.invoice\_id = cte.invoice\_id

left JOIN track t on t.track\_id = il.track\_id

left JOIN album al on al.album\_id = t.album\_id

GROUP BY 1,2,3;

WITH cte as

(

SELECT invoice\_id, COUNT(DISTINCT a.name) num FROM invoice\_line il

left JOIN track t on t.track\_id = il.track\_id

left JOIN album al on al.album\_id = t.album\_id

left join artist a on a.artist\_id = al.artist\_id

GROUP BY 1 HAVING COUNT(DISTINCT a.name) > 1

)

SELECT cte.invoice\_id, num, a.name FROM cte

left join invoice\_line il on il.invoice\_id = cte.invoice\_id

left JOIN track t on t.track\_id = il.track\_id

left JOIN album al on al.album\_id = t.album\_id

left join artist a on a.artist\_id = al.artist\_id

GROUP BY 1,2,3;

OUTPUT :



Insights : When the output of the above query is plotted as a table in excel and a pivot table is constructed with genres in rows, invoice id in columns and count of genres in values it can be seen that the

* Rock, Metal & Alternative and Punk are the genres frequently purchased together.

In similar way when when albums are checked then

* Mezmerize, The Doors & Dark Side Of The Moon are the albums frequently purchased together.

In simalar way when artists are checked then

* Green Day, Foo Fighters & U2 are the artists frequently purchased together.

Recommendations:

* Use these insights for cross-selling by bundling related products.
* Highlight popular combinations in marketing campaigns to drive more sales.

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?**

SQL Query: -

WITH num\_cust\_in\_1st\_3months as

(

SELECT billing\_country, COUNT(customer\_id) ttl from invoice

WHERE invoice\_date BETWEEN '2017-01-01' AND '2017-03-31'

GROUP BY 1

),

num\_cust\_in\_last\_2months as

(

SELECT billing\_country, COUNT(customer\_id) l\_num FROM invoice

WHERE invoice\_date BETWEEN '2020-11-01' AND '2020-12-31'

GROUP BY 1

)

SELECT n1.billing\_country, (ttl - COALESCE(l\_num,0))/ttl \* 100 churn\_rate FROM num\_cust\_in\_1st\_3months n1

LEFT JOIN num\_cust\_in\_last\_2months n2 on n1.billing\_country = n2.billing\_country;

WITH num\_cust\_in\_1st\_3months as

(

SELECT billing\_city, COUNT(customer\_id) ttl from invoice

WHERE invoice\_date BETWEEN '2017-01-01' AND '2017-03-31'

GROUP BY 1

),

num\_cust\_in\_last\_2months as

(

SELECT billing\_city, COUNT(customer\_id) l\_num FROM invoice

WHERE invoice\_date BETWEEN '2020-11-01' AND '2020-12-31'

GROUP BY 1

)

SELECT n1.billing\_city, (ttl - COALESCE(l\_num,0))/ttl \* 100 churn\_rate FROM num\_cust\_in\_1st\_3months n1

LEFT JOIN num\_cust\_in\_last\_2months n2 on n1.billing\_city = n2.billing\_city;

WITH num\_cust\_in\_1st\_3months as

(

SELECT billing\_state, COUNT(customer\_id) ttl from invoice

WHERE invoice\_date BETWEEN '2017-01-01' AND '2017-03-31'

GROUP BY 1

),

num\_cust\_in\_last\_2months as

(

SELECT billing\_state, COUNT(customer\_id) l\_num FROM invoice

WHERE invoice\_date BETWEEN '2020-11-01' AND '2020-12-31'

GROUP BY 1

)

SELECT n1.billing\_state, (ttl - COALESCE(l\_num,0))/ttl \* 100 churn\_rate FROM num\_cust\_in\_1st\_3months n1

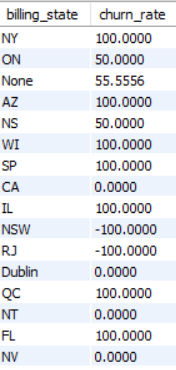
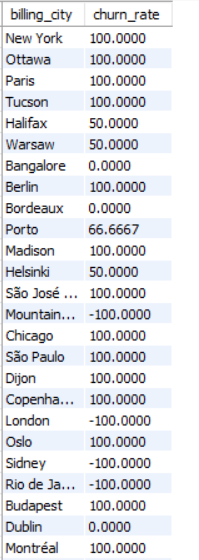
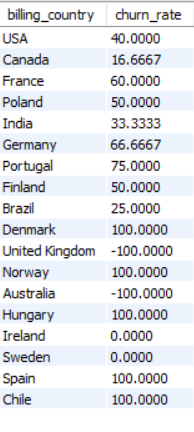
LEFT JOIN num\_cust\_in\_last\_2months n2 on n1.billing\_state = n2.billing\_state;

SELECT billing\_country, COUNT(invoice\_id) num\_invoices, AVG(total) avg\_sales FROM invoice

GROUP BY 1

ORDER BY COUNT(invoice\_id) DESC, AVG(total) DESC

OUTPUT:



Insights: - Therefore it can be seen from the above queries that the customer churn rate varies across different countries, cities and state. It is evident that the European countries have a higher churn rate compared to other countries. And yes, the purchasing behaviour of the customers vary across different countries.

Recommendations: - It can be observed that the developed nations have more number of orders and average sales as compared to the developing nations. It shows that economic factors play a crucial role in the sales generated from a region. Therefore, a greater number of advertisements should be launched in the countries with high economy and cheaper tracks should be sold in the countries with low economy. Population of countries can also be considered as key metric in the analysis.

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?**

SQL Query: -

SELECT i.customer\_id, CONCAT(first\_name, " ", last\_name) name, billing\_country, invoice\_date, SUM(total) total\_spending, COUNT(invoice\_id) num\_of\_orders FROM invoice i

LEFT JOIN customer c on c.customer\_id = i.customer\_id

GROUP BY 1,2,3,4

ORDER BY name

OUTPUT:

A screenshot of a computer screen

Description automatically generated

Insights: After analyzing the data in the form of charts and tables it can be seen that the countries with already high spending amount and frequency of orders, their numbers are increasing whereas the sales and frequency are stagnant in other countries. Therefore, it is seen that new promotional campaigns need to be done in those countries to reduce churn rate as well as maintain & increase the spending.

Factors which contribute to this risk are: -

* Are younger customers more likely to churn? (If information was given)
* Does gender play a role? (if information was given)
* How does location impact churn?
* Analysis of spending behaviour (e.g., high spenders vs. infrequent buyers).

If there was information regarding age and gender of the customers, the customer segmentation would have been –

* Young-Male-High-Spendors
* Young-Female-High-Spendors
* Old-Male-Low-Spendors
* Old-Female-Low-Spendors

and many more.

1. **Customer Lifetime Value Modelling: 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?**

SQL Query :

WITH CustomerTenure AS (

SELECT

c.customer\_id, CONCAT(c.first\_name,' ', c.last\_name) AS customer,

MIN(i.invoice\_date) AS first\_purchase\_date,

MAX(i.invoice\_date) AS last\_purchase\_date,

DATEDIFF(MAX(i.invoice\_date), MIN(i.invoice\_date)) AS tenure\_days,

COUNT(i.invoice\_id) AS purchase\_frequency,

SUM(i.total) AS total\_spent

FROM customer c

JOIN invoice i ON c.customer\_id = i.customer\_id

GROUP BY c.customer\_id

)

SELECT

customer\_id,

customer,

tenure\_days,

purchase\_frequency,

total\_spent,

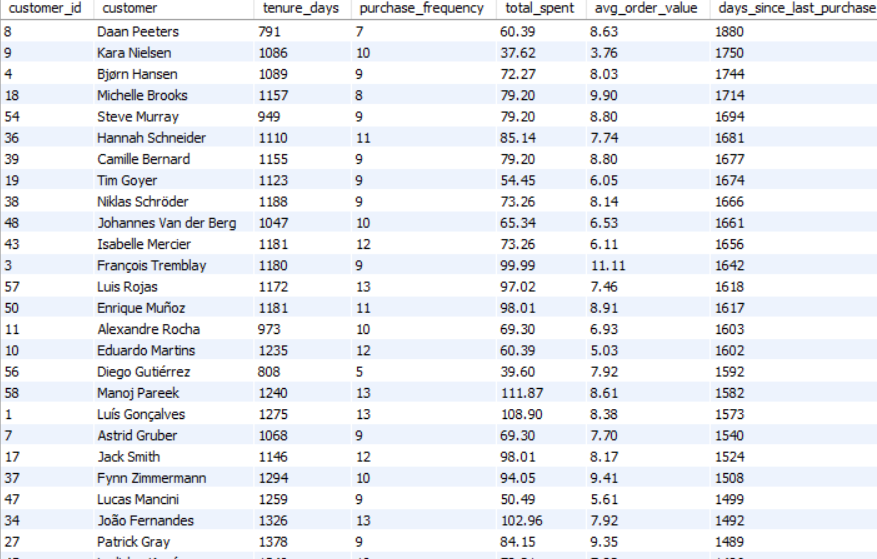
ROUND(total\_spent / purchase\_frequency, 2) AS avg\_order\_value,

DATEDIFF(CURRENT\_DATE, last\_purchase\_date) AS days\_since\_last\_purchase

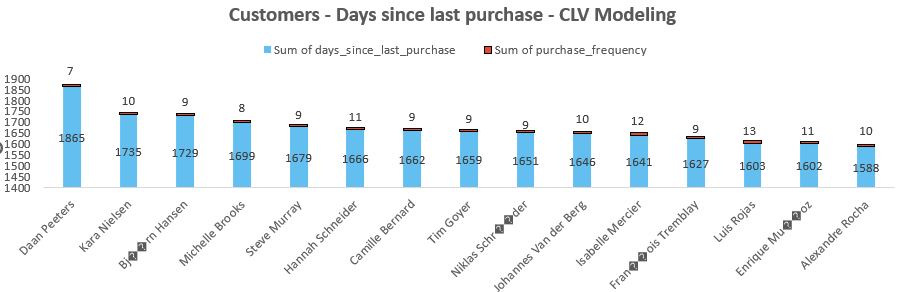
FROM CustomerTenure

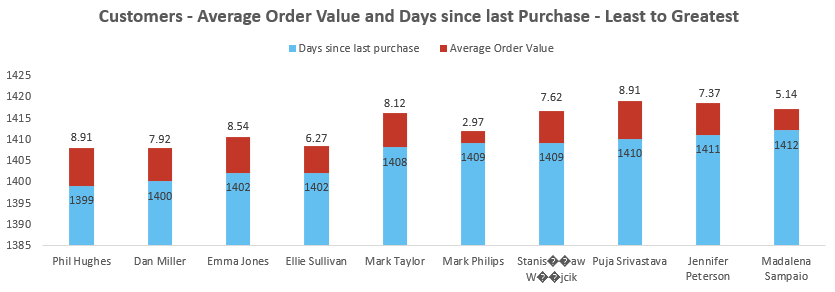
ORDER BY days\_since\_last\_purchase DESC;

OUTPUT:



* 1. To predict the lifetime value of the customer segment, I could analyse the customer purchase history and the tenure for which a customer is with the company. With these two parameters we can judge whether a customer is a High\_value customer or not.
  2. If the tenure of the customer is small but the purchase history is big, then promotional campaigns should be targeted to these type of customers as they can be converted into loyal customers.
  3. The customers who have stopped purchasing have one thing in common that they belong to the under developed or developing countries. This shows that economic factor plays an important role in the sales of the company. This could be prevented by using adequate number of promotion channels like social media, articles and advertisements. Also giving discounts could boost the sales in those countries.





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?**

If the data on promotional campaigns was available, I would have used it to analyze its impact on

1. Customer Acquisition based on

* The increase in the number of customers with time.
* Number of people participating in the events held in different locations.
* The increase in click-through rate due to the email marketing campaigns.

1. Retention on the basis of

* Number of old customers attending the event and then purchasing the track again after long time.
* Number of customers getting a discount.

1. Sales on the basis of

* Increase in the sales due to promotional campaigns.
* Trends to analyse which promotional campaign was the best.
* New customers being generated due to discounts.

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

If the objective and subjective questions weren’t given, I would have first gone through the aim of the project and based on that decide the KPIs which would have helped me reach the objective.

Then I would have developed various business questions like: -

* Which are locations from where maximum sale is coming?
* Which tracks are the best-selling in each location?
* Which are the most popular artists and albums in different regions?
* How is sales distributed among different countries?
* Which genre is most liked in each country?

& also look over the Trends in the purchases of tracks like: -

* What is the total revenue generated over a period of time?
* How is the invoice total of each customer changed with time?
* What is the trend of a particular genre in a particular country?

Then I would have summarized all the insights and presented the data.

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?**

We can use the ALTER TABLE command to add a new column to the Album table.

SQL Query:

ALTER TABLE album

ADD COLUMN ReleaseYear INT(4);

SELECT \* FROM album;data\_type int;

OUTPUT :



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.**

SQL Query: -

SELECT billing\_country,

COUNT(DISTINCT customer\_id) num\_of\_customers,

AVG(total) Average\_total\_amount,

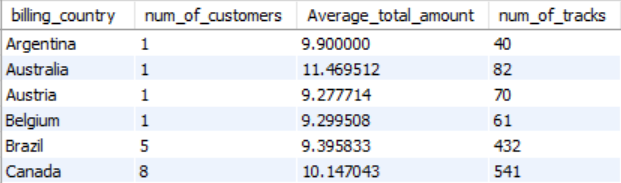
COUNT(track\_id) num\_of\_tracks

FROM invoice i

LEFT JOIN invoice\_line il on il.invoice\_id = i.invoice\_id

GROUP BY 1 ;

OUTPUT :

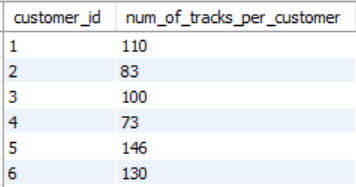


SELECT customer\_id, COUNT(DISTINCT track\_id) num\_of\_tracks\_per\_customer FROM invoice i

LEFT JOIN invoice\_line il on il.invoice\_id = i.invoice\_id

GROUP BY 1

OUTPUT: -



---------------------THE END-----------------------