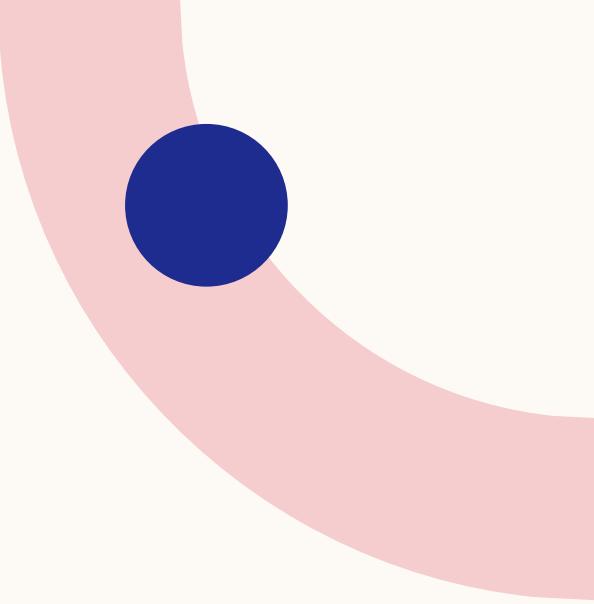
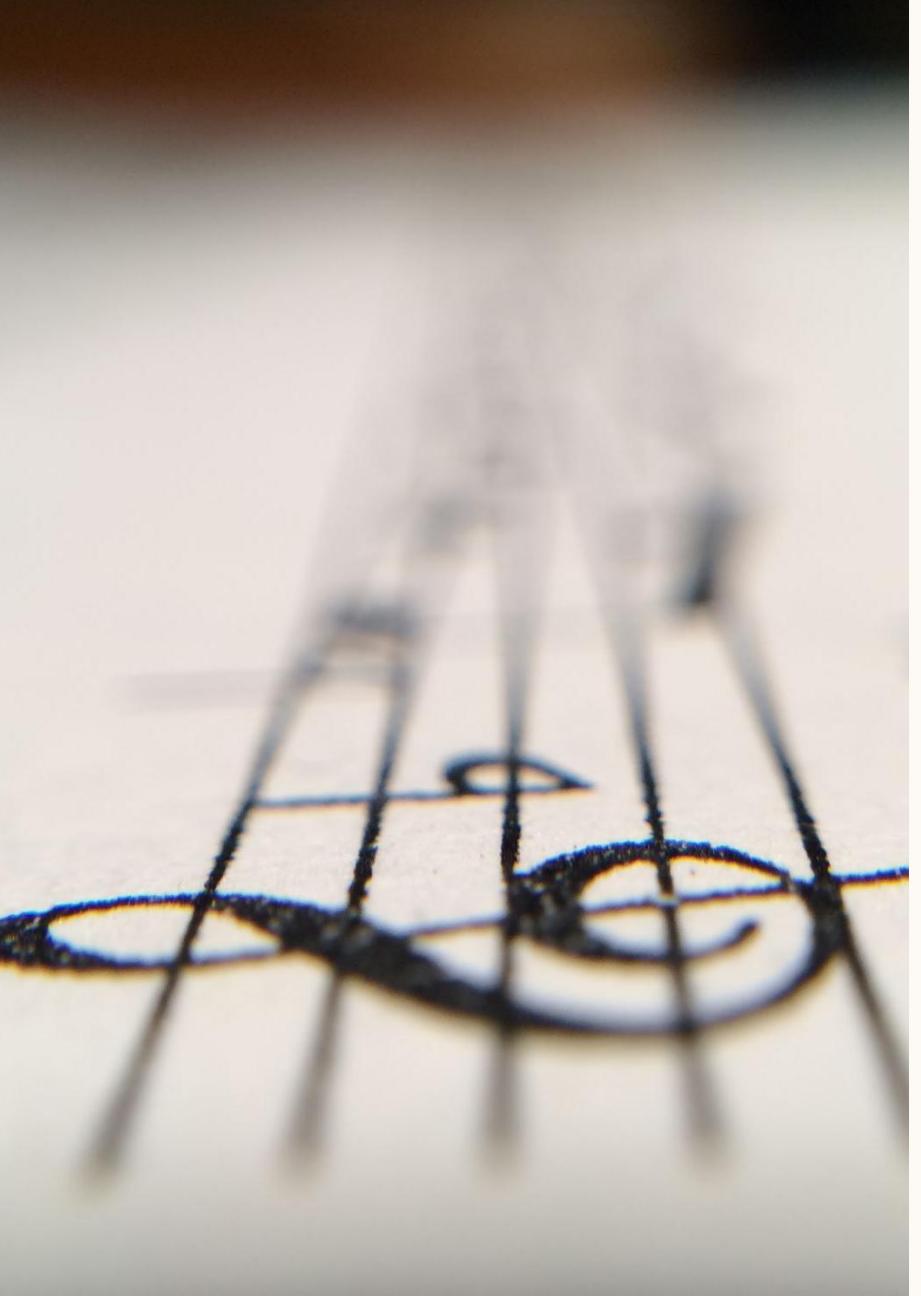


CHINOOK SQL PROJECT

**BY
SNEHHA WANKHEDE**



CUSTOMER ANALYSIS & INSIGHTS FROM CHINOOK DATASET

AGENDA

Problem Statement

Data Description

Objective Key Metrics and Visualizations

Subjective Question for Insights

Final tips & takeaways

PROBLEM STATEMENT

- Chinook faces inconsistent retention across geographies
- Understand customer churn behavior
- Identify high-value customers and lifetime value patterns
- Profile customer segments based on geography and purchase behavior
- Generate actionable recommendations to reduce churn and boost revenue

DATA DESCRIPTION

Data Summary:

11 relational tables

1. Album (347 records)

2. Artist (275 records)

3. Customer (59 records)

4. Employee (8 records)

5. Genre (25 records)

6. Invoice (614 records)

7. Invoice_line (1000 records)

8. Media_type (5 records)

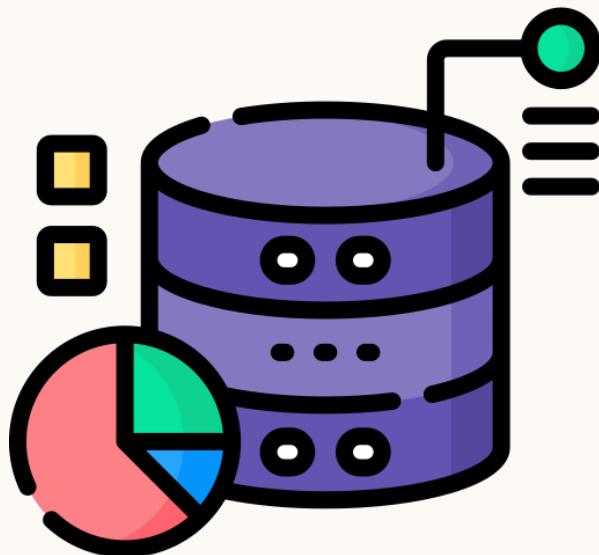
9. Playlist (18 records)

10. Playlist_track (1000 records)

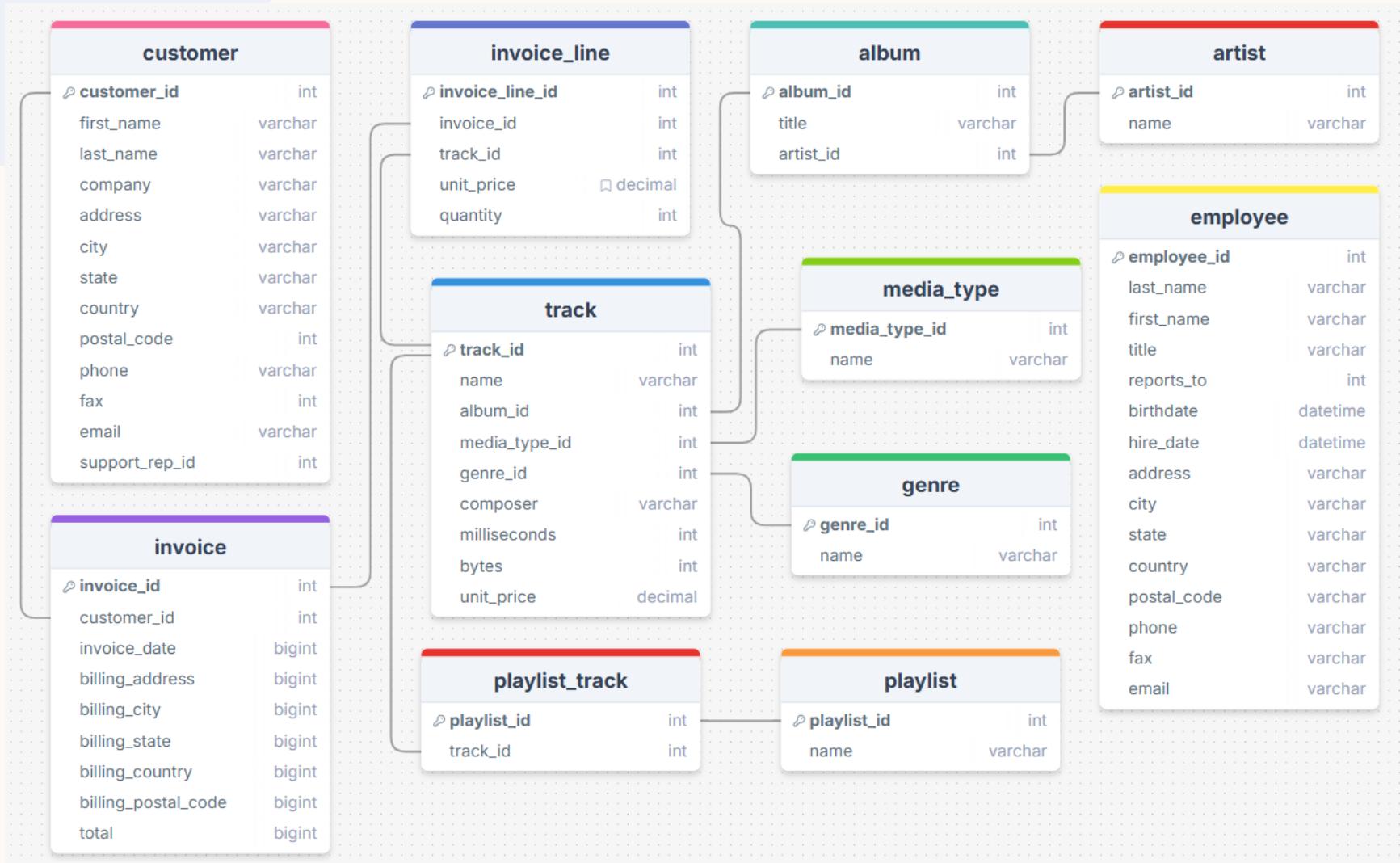
11. Track (1000 records)

Preprocessing

- Identify NULL & duplicate records
- Normalized date formats



DATABASE SCHEMA



METHODOLOGY

1. SQL (MySQL)
2. Excel (for visualizations)

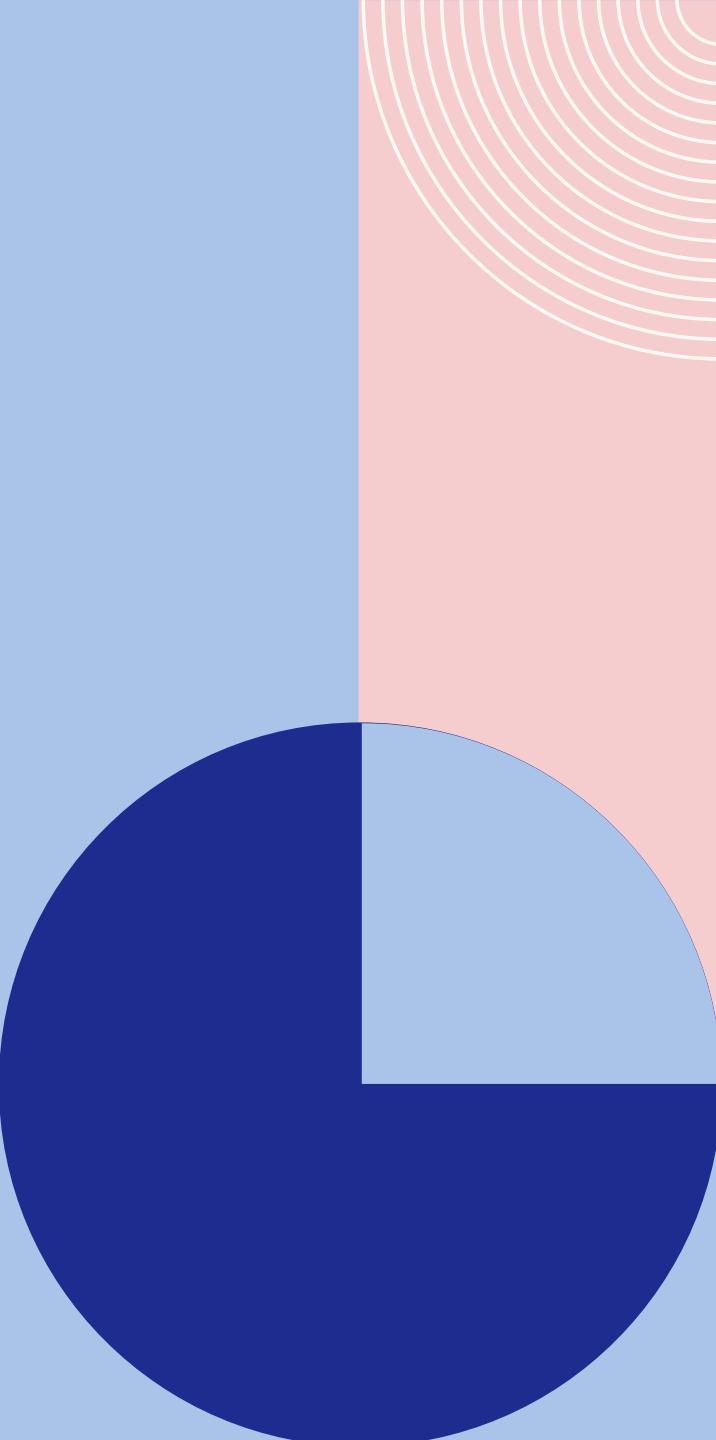
Steps Taken:

1. Data extraction & joins from Chinook tables
2. Created metrics: Total Spent, Days Since Last Purchase, Risk Levels, Churn Rate
3. Segmented customers based on geography and purchasing patterns

Key Functions:

1. DATEDIFF(), DATE_SUB(), GROUP BY, ORDER BY, JOINS, CASE WHEN, etc. to execute the desired output





OBJECTIVE KEY METRICS AND VISUALIZATIONS

Chinook Sample Database

1. MISSING VALUES / DUPLICATES

Key Metric: Identify the missing values, NULL or duplicates

Visualization: Table

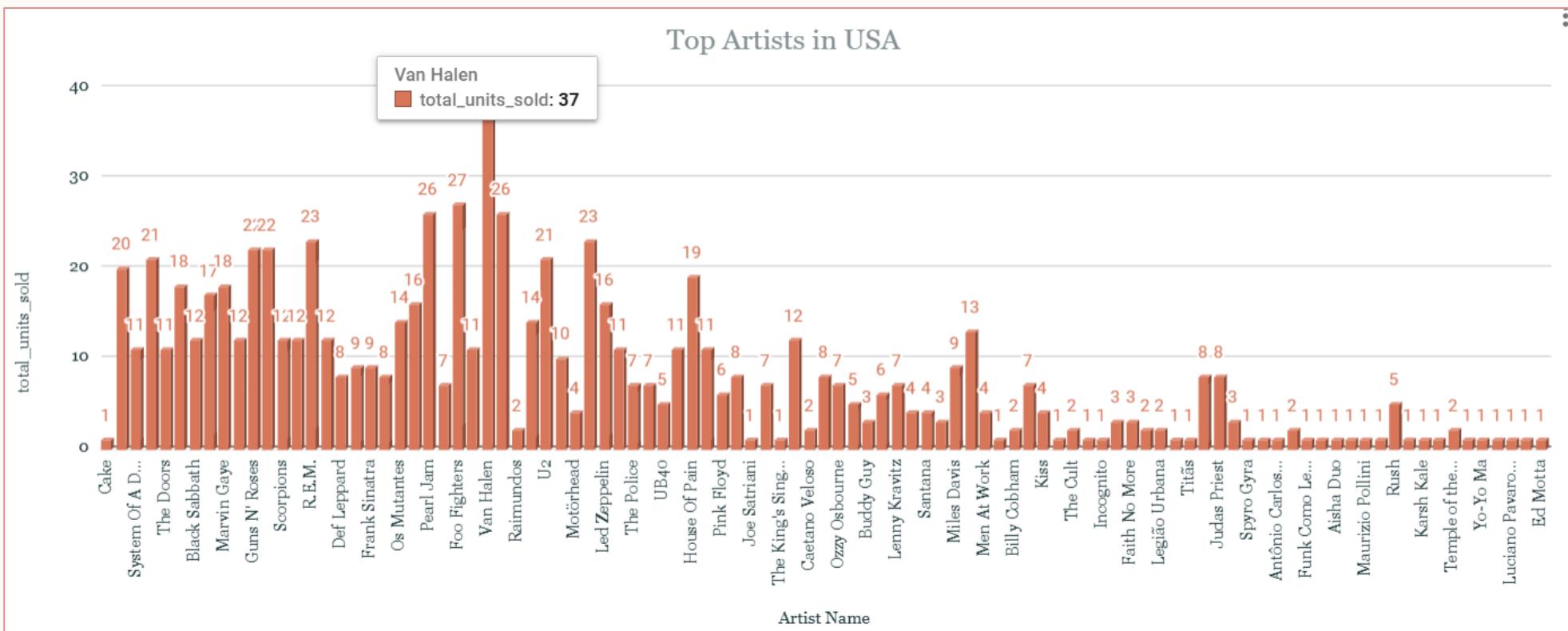
Table Name	Data Validation
Customer	<ul style="list-style-type: none">• 49 records are missing in Company column• 29 records are missing in State column• 4 postal code records are missing• 1 phone missing• 47 fax values are missing
Employee	1 record is NULL/missing in reports_to column
Track	978 records are missing in Composer column
Invoice	304 records mentioned None in billing_state column

2. TOP TRACKS/ARTISTS IN USA

Key Metric: Total items sold per Artist/track, Genre-wise distribution

Visualization: Column Chart

Insight: Artist Van Halen is the top artist in UAS



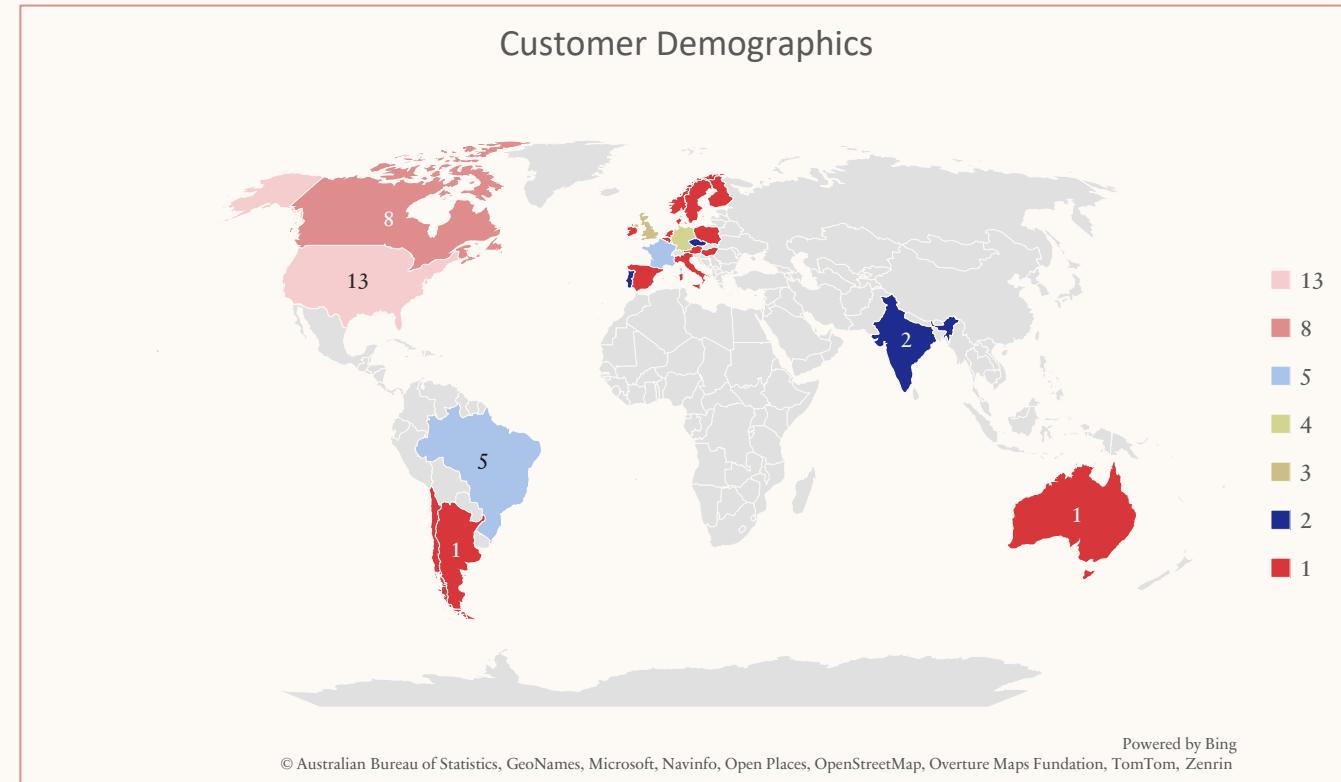
3. CUSTOMER DEMOGRAPHICS

Key Metric: Count by Country (Gender, Age Group data is not provided)

Visualization: Query output snapshot, Filled Map for Geo Distribution

Insight: Among the 24 countries USA have the highest customer base

country	customer_count
USA	13
Canada	8
Brazil	5
France	5
Germany	4
United Kingdom	3
Czech Republic	2
Portugal	2
India	2
Norway	1
Austria	1
Belgium	1
Denmark	1
Finland	1
Hungary	1
Ireland	1
Italy	1
Netherlands	1
Poland	1
Spain	1
Sweden	1
Australia	1
Argentina	1
Chile	1



4. REVENUE FOR EACH COUNTRY/STATE/CITY

Key Metric: Find out the revenue for each country, state, and city

Visualization: Query output snapshot

Insight: Among countries, the USA generates the highest revenue. At the state level, entries marked as 'None' account for the top revenue figures. Among cities, Prague leads in revenue generation

Country	TotalInvoices	TotalRevenue
USA	131	1040.49
Canada	76	535.59
Brazil	61	427.68
France	50	389.07
Germany	41	334.62
Czech Republic	30	273.24
United Kingdom	28	245.52
Portugal	29	185.13
India	21	183.15
Ireland	13	114.84
Spain	11	98.01
Chile	13	97.02
Australia	10	81.18
Finland	11	79.20
Hungary	10	78.21
Poland	10	76.23
Sweden	10	75.24
Norway	9	72.27
Austria	9	69.30
Netherlands	10	65.34
Belgium	7	60.39
Italy	9	50.49
Argentina	5	39.60
Denmark	10	37.62

City	TotalInvoices	TotalRevenue
Prague	30	273.24
Mountain View	20	169.29
London	19	166.32
Berlin	20	158.40
Paris	18	151.47
São Paulo	22	129.69
Dublin	13	114.84
Delhi	13	111.87
São José dos Campos	13	108.90
Brasília	15	106.92
Lisbon	13	102.96
Bordeaux	11	99.99
Montréal	9	99.99
Madrid	11	98.01
Redmond	12	98.01
Santiago	13	97.02
Frankfurt	10	94.05
Orlando	12	92.07
Ottawa	13	91.08
Reno	11	91.08
Fort Worth	12	86.13
Tucson	9	84.15
Porto	16	82.17
Stuttgart	11	82.17
Rio de Janeiro	11	82.17

State	TotalInvoices	TotalRevenue
None	304	2393.82
SP	35	238.59
CA	29	223.74
ON	17	131.67
Dublin	13	114.84
DF	15	106.92
QC	9	99.99
WA	12	98.01
FL	12	92.07
NV	11	91.08
TX	12	86.13
AZ	9	84.15
RJ	11	82.17
NSW	10	81.18
NY	8	79.20
WI	10	76.23
NT	12	75.24
UT	10	72.27
IL	8	71.28
MB	8	70.29
BC	9	66.33
MA	10	66.33
VV	10	65.34
NS	11	62.37
RM	9	50.49

5. TOP 5 CUSTOMERS PER COUNTRY

Key Metric: Customer-wise revenue ranking

Visualization: Query output snapshot

Insight:

- The highest revenue-generating customers are spread across multiple countries, with Brazil, USA, Germany, France, and Canada contributing significantly.
- In USA, the top 5 customers together spent over 463, showing strong market engagement.
- This suggests that personalized loyalty campaigns could be designed for high-revenue countries, while smaller markets with high-value individuals (e.g. Czech Republic, Ireland) may benefit from targeted premium offerings.

CustomerId	FirstName	LastName	BillingCountry	Total_Revenue	TotalByCountry
56	Diego	Gutiérrez	Argentina	39.6	39.6
55	Mark	Taylor	Australia	81.18	81.18
7	Astrid	Gruber	Austria	69.3	69.3
8	Daan	Peeters	Belgium	60.39	60.39
1	Luis	Gonçalves	Brazil	108.9	
13	Fernanda	Ramos	Brazil	106.92	
12	Roberto	Almeida	Brazil	82.17	
11	Alexandre	Rocha	Brazil	69.3	
10	Eduardo	Martins	Brazil	60.39	
3	François	Tremblay	Canada	99.99	
30	Edward	Francis	Canada	91.08	
33	Ellie	Sullivan	Canada	75.24	
32	Aaron	Mitchell	Canada	70.29	
15	Jennifer	Peterson	Canada	66.33	
57	Luis	Rojas	Chile	97.02	
5	František	Wichterlová	Czech Republic	144.54	
6	Helena	Holý	Czech Republic	128.7	
9	Kara	Nielsen	Denmark	37.62	
44	Terhi	Hämäläinen	Finland	79.2	
42	Wyatt	Girard	France	99.99	
39	Camille	Bernard	France	79.2	
43	Isabelle	Mercier	France	73.26	
40	Dominique	Lefebvre	France	72.27	
41	Marc	Dubois	France	64.35	
37	Fynn	Zimmermann	Germany	94.05	
36	Hannah	Schneider	Germany	85.14	
2	Leonie	Köhler	Germany	82.17	
38	Niklas	Schröder	Germany	73.26	
45	Ladislav	Kovács	Hungary	78.21	
58	Manoj	Pareek	India	111.87	
59	Puja	Srivastava	India	71.28	
46	Hugh	O'Reilly	Ireland	114.84	
47	Lucas	Mancini	Italy	50.49	
48	Johannes	Van der Berg	Netherlands	65.34	
4	Bjørn	Hansen	Norway	72.27	
49	Stanislaw	Wójcik	Poland	76.23	
34	João	Fernandes	Portugal	102.96	
35	Madalena	Sampaio	Portugal	82.17	
50	Enrique	Muñoz	Spain	98.01	
51	Joakim	Johansson	Sweden	75.24	
53	Phil	Hughes	United Kingdom	98.01	
54	Steve	Murray	United Kingdom	79.2	
52	Emma	Jones	United Kingdom	68.31	
17	Jack	Smith	USA	98.01	
20	Dan	Miller	USA	95.04	
22	Heather	Leacock	USA	92.07	
21	Kathy	Chase	USA	91.08	
26	Richard	Cunningham	USA	86.13	

6. TOP-SELLING TRACK PER CUSTOMER

14

Key Metric: Identify the top selling track or max revenue for each customer

Visualization: Query output snapshot

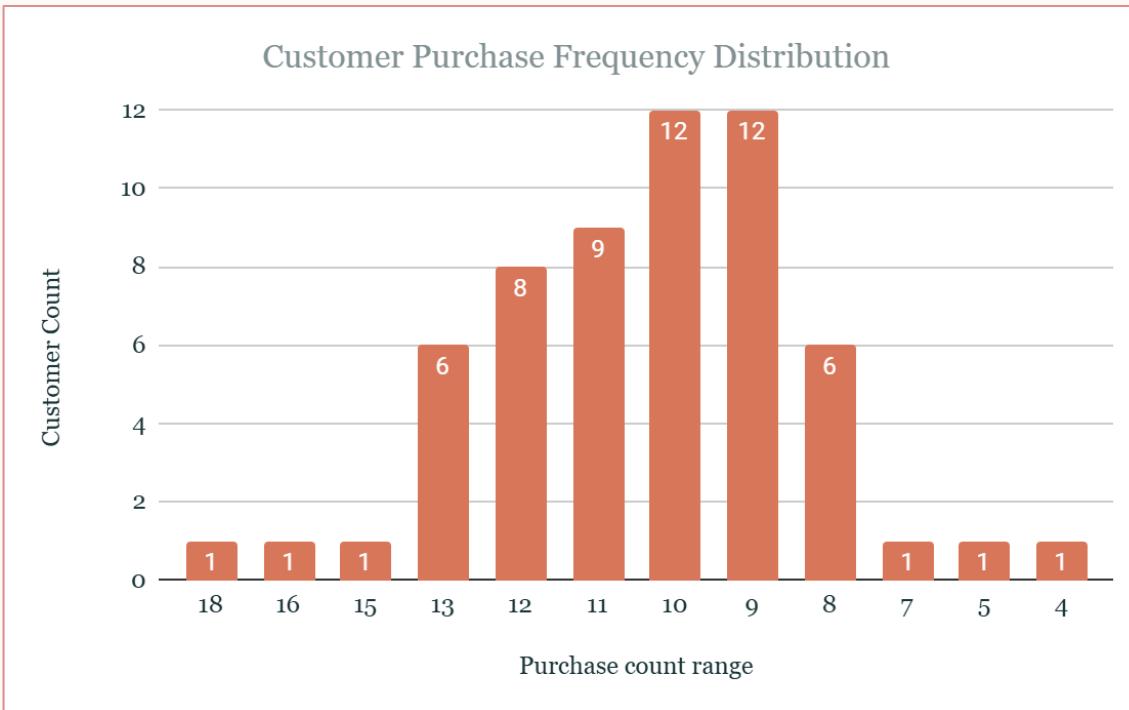
Insight:

- The data shows that for most customers, the top-selling track is purchased only once or twice
- Diverse preferences across countries and customers
- High-revenue tracks are generated above-average revenue (19.80 - 21.78)

customer_id	first_name	last_name	track_name	total_units	total_revenue
12	Roberto	Almeida	Love And Marriage	2	21.78
13	Fernanda	Ramos	24 Caprices, Op. 1, No. 24, for Solo Vi...	2	19.80
3	François	Tremblay	Sting Me	2	17.82
28	Julia	Barnett	Get What You Need	2	19.80
10	Eduardo	Martins	Like A Bird	2	9.90
37	Fynn	Zimmermann	Radio/Video	2	9.90
34	João	Fernandes	Train In Vain	2	19.80
43	Isabelle	Mercier	Tease Me Please Me	2	19.80
46	Hugh	O'Reilly	N.I.B.	2	19.80
48	Johannes	Van der Berg	Confusion	2	15.84
49	Stanisław	Wójcik	Faceless	2	17.82
7	Astrid	Gruber	Saint Of Me	1	9.90
8	Daan	Peeters	I Wish It Would Rain	1	5.94
9	Kara	Nielsen	Send Me an Angel	1	0.99
11	Alexandre	Rocha	That's Life	1	1.98
14	Mark	Philips	Us And Them	1	1.98
15	Jennifer	Peterson	Walking Contradiction	1	7.92
16	Frank	Harris	Eruption	1	16.83
17	Jack	Smith	Scentless Apprentice	1	9.90
18	Michelle	Brooks	Balada Do Louco	1	13.86
19	Tim	Goyer	I Looked At You	1	8.91
20	Dan	Miller	Still	1	9.90
21	Kathy	Chase	Where Have All The Good Times Gone?	1	11.88
22	Heather	Leacock	I'm A Boy	1	8.91
24	Frank	Ralston	War Pigs	1	9.90

7. PURCHASE BEHAVIOR TRENDS

15



Key Metric: Frequency, Avg Order Value

Insight:

- The graph shows a bell-curve-like pattern, suggesting normal buying behavior
- The peak is at 9 and 10 purchases, with 12 customers each

Avg Order Value
\$11.11

- Followed by 11 (9 customers) and 12 (8 customers)
- Very few customers make extremely high (15–18) or very low (4–5) purchases.
- A small group of loyal customers made 16–18 purchases
- Customers with ≤ 7 purchases may need re-engagement campaigns

8. CUSTOMER CHURN RATE

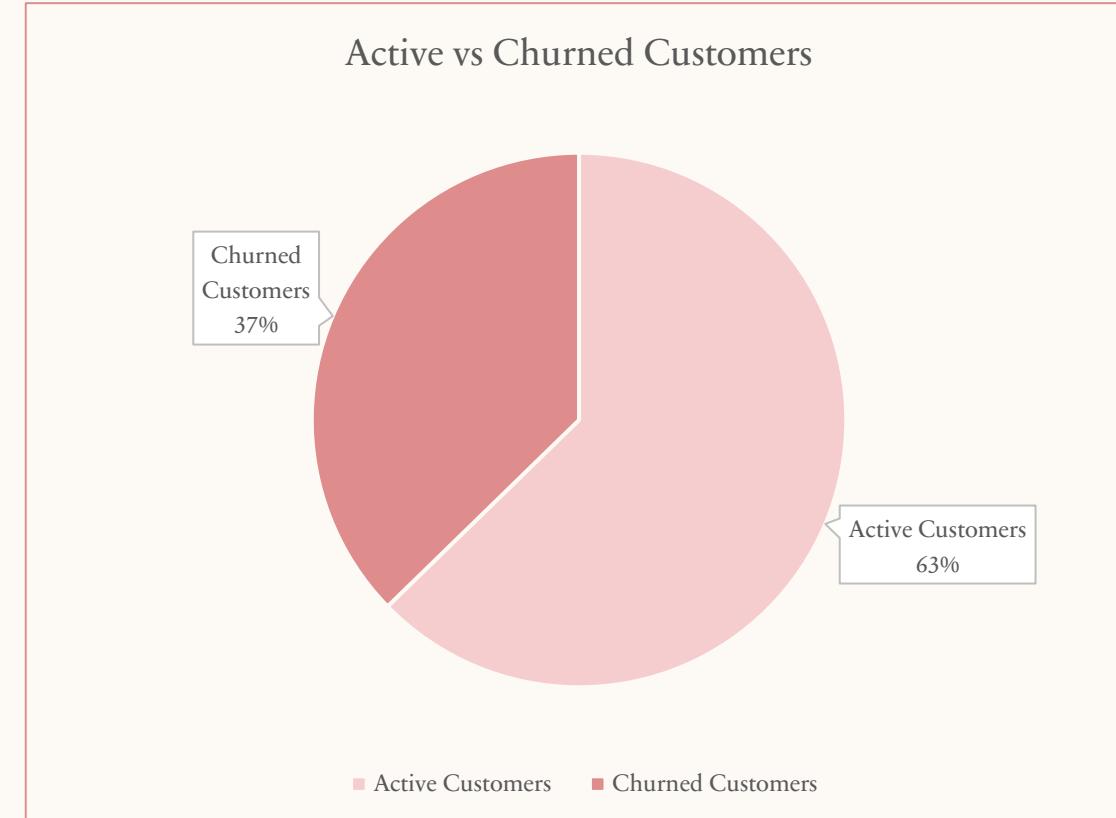
Key Metric: Churn rate

Insight:

Approximately 37% customers has not made any purchases in the last 3 months, indicating a significant customer churn.

Calculation-

- Total Customers = 59
- Churned Customers = 22 (inactive since last 3 months)
- Churn Rate = $(15 / 59) \times 100 \approx 37.29\%$



9. SALES BY GENRE IN USA

17

Key Metric: Genre-wise sales contribution and top artist in USA

Insight:

- Rock dominates USA sales at 53%, led by Van Halen.
- Genres like Alternative & Punk 12% and Metal 11% show some listener base
- Artists like Green Day, Black Sabbath are the top performer followed by Van Halen
- Lower engagement in Jazz, Hip Hop, Classical need a growth or targeted campaigns.

GenreName	TotalSales	PercentOfTotal(%)	TopArtist
Rock	555.39	53.38	Van Halen
Alternative & Punk	128.7	12.37	Green Day
Metal	122.76	11.8	Black Sabbath
R&B/Soul	52.47	5.04	Marvin Gaye
Blues	35.64	3.43	Eric Clapton
Alternative	34.65	3.33	Chris Cornell
Latin	21.78	2.09	Cássia Eller
Pop	21.78	2.09	Amy Winehouse
Hip Hop/Rap	19.8	1.9	House Of Pain
Jazz	13.86	1.33	Miles Davis
Easy Listening	12.87	1.24	Frank Sinatra
Reggae	5.94	0.57	UB40
Electronica/Dance	4.95	0.48	Jamiroquai
Classical	3.96	0.38	The King's Singers
Heavy Metal	2.97	0.29	Iron Maiden
Soundtrack	1.98	0.19	Gilberto Gil
TV Shows	0.99	0.1	Lost

11. GENRE RANKING IN USA BASED ON SALES PERFORMANCE

Key Metric: Genre sales performance in the USA

Insight:

- Top 3 genres Rock, Alternative & Punk, Metal drive majority of revenue
- Rock ranking on first position with 555.39 sales indicating a strong preference among U.A listeners.
- Combined sales of Rock, Alternative & Punk, and Metal account for over 77% of total revenue
- R&B/Soul, Blues, and Alternative show moderate performance
- Genres like TV Shows, Soundtrack, Heavy Metal, and Classical contribute less revenue each, these may either have to promote more, or deprioritize

Genre_name	Total_sales	Sales_rank
Rock	555.39	1
Alternative & Punk	128.70	2
Metal	122.76	3
R&B/Soul	52.47	4
Blues	35.64	5
Alternative	34.65	6
Latin	21.78	7
Pop	21.78	7
Hip Hop/Rap	19.80	9
Jazz	13.86	10
Easy Listening	12.87	11
Reggae	5.94	12
Electronica/Dance	4.95	13
Classical	3.96	14
Heavy Metal	2.97	15
Soundtrack	1.98	16
TV Shows	0.99	17

12. INACTIVE CUSTOMERS (LAST 3 MONTHS)

customer_id	first_name	last_name	LastPurchaseDate
18	Michelle	Brooks	2020-03-05
38	Niklas	Schröder	2020-04-22
1	Luís	Gonçalves	2020-07-24
10	Eduardo	Martins	2020-06-25
43	Isabelle	Mercier	2020-05-02
9	Kara	Nielsen	2020-01-29
4	Bjørn	Hansen	2020-02-04
39	Camille	Bernard	2020-04-11
3	François	Tremblay	2020-05-16
58	Manoj	Pareek	2020-07-15
37	Fynn	Zimmermann	2020-09-27
50	Enrique	Muñoz	2020-06-10
19	Tim	Goyer	2020-04-14
36	Hannah	Schneider	2020-04-07
57	Luis	Rojas	2020-06-09
48	Johannes	Van der Berg	2020-04-27
8	Daan	Peeters	2019-09-21
17	Jack	Smith	2020-09-11
54	Steve	Murray	2020-03-25
7	Astrid	Gruber	2020-08-26
11	Alexandre	Rocha	2020-06-24
56	Diego	Gutiérrez	2020-07-05

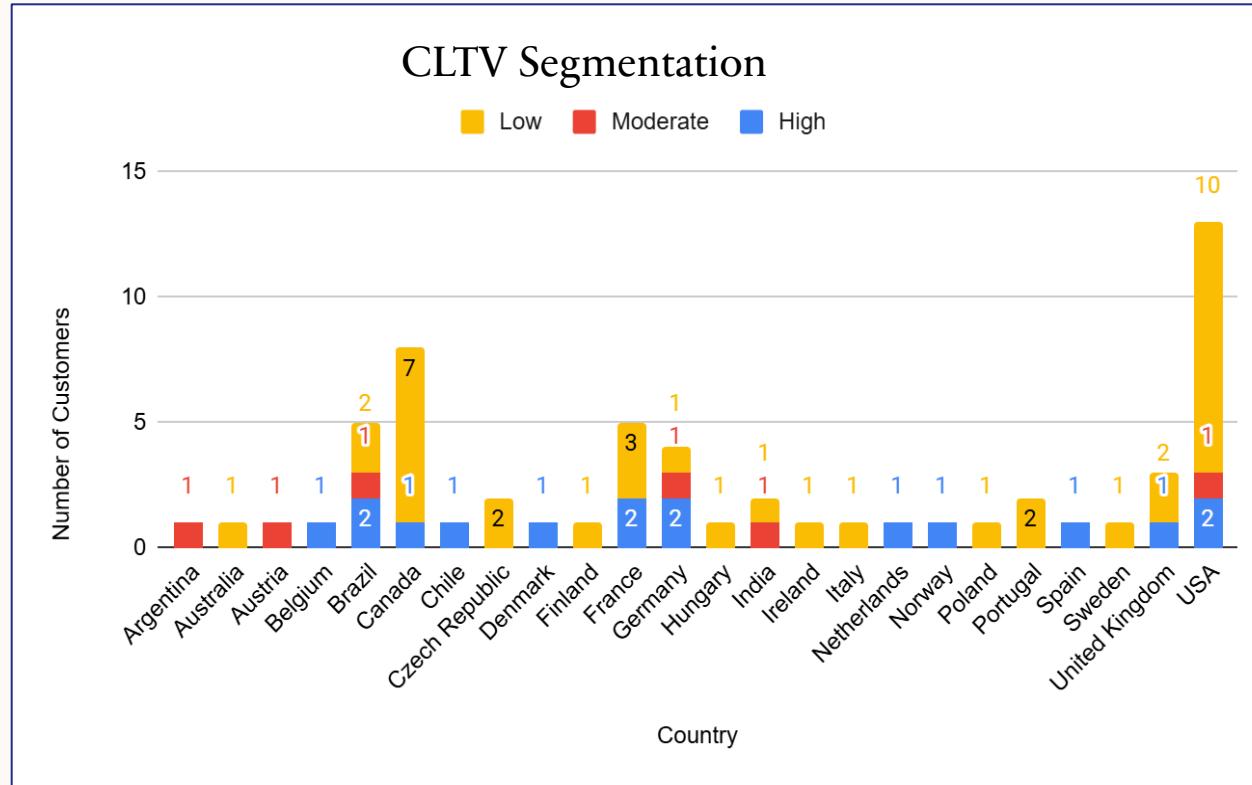
Key Metric: Inactive customers list

Insight: Total inactive customers 22

- These customers haven't made a purchase since before 30th September 2020.
- The longest inactive customer is Daan Peeters, last purchase on 2019-09-21.
- A few customers, like Fynn Zimmermann (2020-09-27) and Luis Goncalves (2020-07-24), were active in mid to late 2020 but have since churned.
- These customers could be **at risk of permanent churn** without re-engagement strategies.

- Need to target these 22 customers with reactivation campaigns, send personalized emails, or discounts.

CLTV SEGMENTATION AND STRATEGIC RECOMMENDATIONS



High CLTV

- Brazil (2), USA (2), France (2), Germany (2), Belgium, Canada, Denmark, Netherlands, Norway, Spain (1 each)
- These customers are driving maximum revenue with high purchase frequency or value

Low CLTV

- USA (10), Canada (7), France (3), Portugal (2), Germany (1), Czech Republic(2), rest each (1)
- Reflects poor engagement or satisfaction

Moderate CLTV

- Argentina, Austria, Brazil, Germany, India, USA shows 1 each
- These are warm prospects who can be encouraged towards higher spending

- Here is the CLTV(Customer life time value) Segmentation by Region chart, showing the distribution of High, Mid, and Low LTV customers across key regions.

STRATEGIC RECOMMENDATIONS

- High CLTV - Retain and Upsell:
 - Focus Countries: Brazil, USA, France, Germany
 - Actions: Loyalty programs, VIP perks, Beta access to new features
- Moderate CLTV – Upgrade and Engage
 - Focus Countries: Germany, Brazil, India, UK
 - Actions: Upselling, educational or promotional content, limited-time offers
- Low CLTV – Reactivate or Cleanse:
 - Focus Countries: USA, Canada, Portugal
 - Actions: Win-back campaigns, feedback collection, segment pruning if unresponsive



FINAL TIPS & TAKEAWAYS

- Promote high-performing genres and artists through curated campaigns
- Deploy retention strategies (email, discounts, recommendations) for mid-CLTV customers
- Use churn prediction models and build marketing funnel to stay ahead of customer attrition
- Segment customers continuously to adapt to changing behaviors.
- Localized content and campaigns can improve sales performance.
- Many customers inactive in the last 3 months - Ideal targets for win-back campaigns and tailored offers.
- Nurture High CLTV customers with loyalty programs and personalized offers.
- Majority of customers fall into Low CLTV which indicates the need for re-engagement or upselling strategies.

**THANK
YOU**