# SQL Case Study:
Digital Music Store

The Music Store sells a wide variety of music, including albums, tracks, and playlists. It also allows customers to create their own playlists. The store has a large number of employees, who work at multiple locations in Canada. The store also has a large number of customers world-wide, who make regular purchases.

Entity Relationship Diagram —
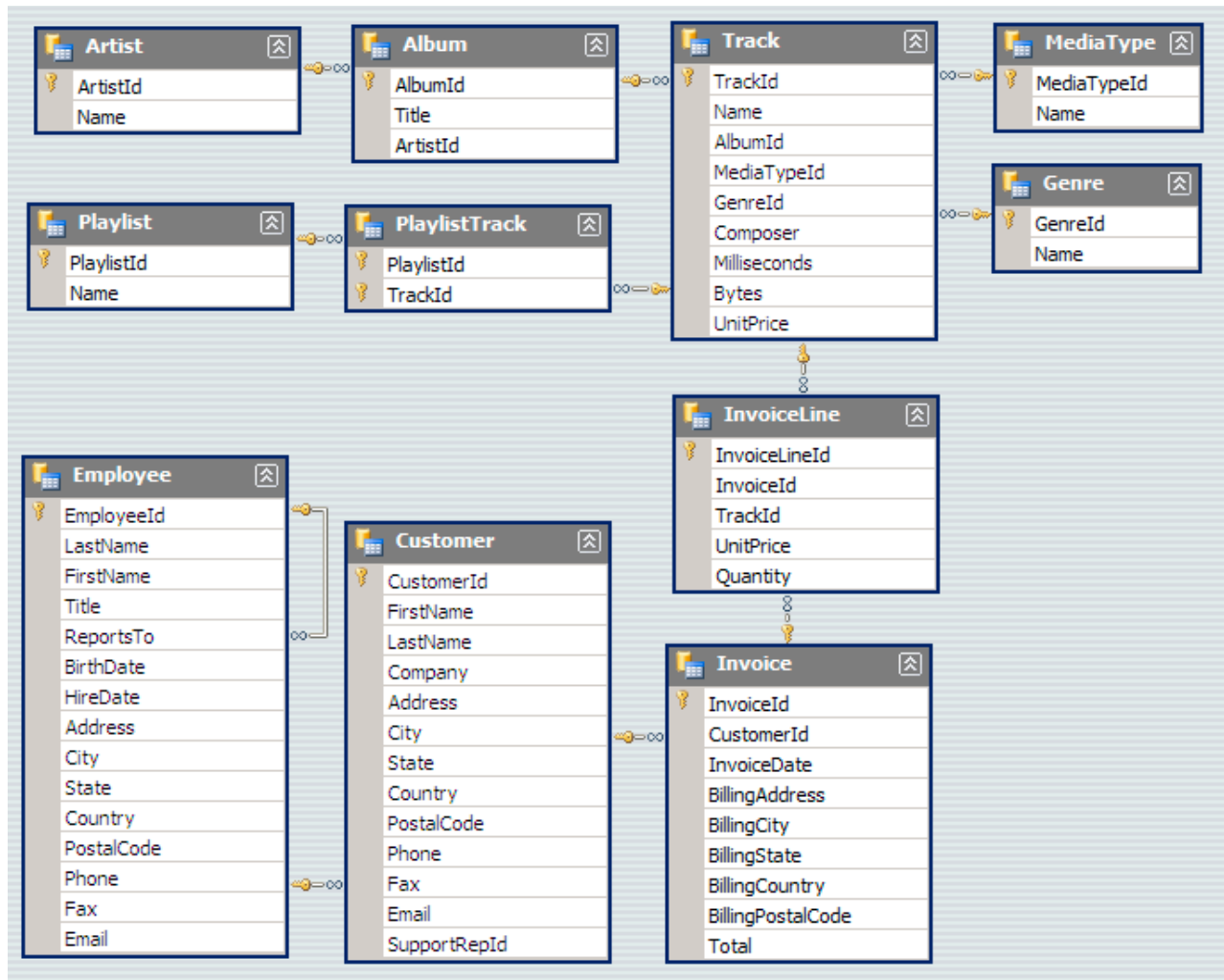
The Data is available at 11 csv tables —

- Artist
- Invoice
- Playlist
- Employee
- Album
- Invoice Line
- Playlist Track
- Genre
- Customer
- Track
- Media Type

Q1. Who is the senior most employee based on job title ?

```sql
SELECT *
FROM employee
ORDER BY levels DESC
LIMIT 1;
```

employee_id	last_name	first_name	title	reports_to	levels	birthdate	hire_date	address	city	state	country	postal_code	phone	fax	email
9	Madan	Mohan	Senior General Manager	NULL	L7	26-01-1961 00:00	14-01-2016 00:00	1008 Vrinda Ave MT	Edmonton	AB	Canada	T5K 2N1	+1 (780) 428-9482	+1 (780) 428-3457	madan.mohan@chinookcorp.com
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Q2. Which countries have the most Invoices ?

```sql
SELECT COUNT(invoice_id) AS invoice, billing_country
FROM invoice
GROUP BY billing_country
ORDER BY invoice DESC;
```

invoice	billing_country
131	USA
76	Canada
61	Brazil
50	France
41	Germany
30	Czech Republic
29	Portugal
28	United Kingdom
21	India
13	Ireland
13	Chile
11	Finland

Insights –

a) The customer base is heavily concentrated in a few countries. The top 3 (USA, Canada, Brazil) account for 268 customers, which is a significant majority of the provided data.

b) The United States is the single largest market by a huge margin (131 customers), followed by Canada and Brazil. This indicates the store's strongest presence is in North and South America.

c) European nations are well-represented but with individually smaller counts. Key markets include France (50), Germany (41), the UK (28), and a cluster of nations with ~10 customers each (e.g., Czech Republic, Portugal, Poland, Sweden).
d) A large number of countries (e.g., India, Chile, Ireland, Finland) have a smaller but notable customer base (10-21 customers). This signifies a widespread international reach beyond the core markets.

Q3. What are the top 3 values of the total invoice ?

```
SELECT *
FROM invoice
ORDER BY total DESC
LIMIT 3;
```

invoice_id	customer_id	invoice_date	billing_address	billing_city	billing_state	billing_country	billing_postal_code	total
183	42	2018-02-09 00:00:00	9, Place Louis Barthou	Bordeaux	None	France	33000	23.759999999999998
92	32	2017-07-02 00:00:00	696 Osborne Street	Winnipeg	MB	Canada	R3L 2B9	19.8
526	5	2020-06-08 00:00:00	Klanova 9/506	Prague	None	Czech Republic	14700	19.8

Q4. Which city has the best customers ? We would like to throw a promotional Music Festival in the city we made the most money. Write a query that returns one city that has the highest sum of invoice totals. Return both the city name & sum of all invoice totals.

```
SELECT billing_city,
       ROUND(SUM(total),2) AS total
FROM invoice
GROUP BY billing_city
ORDER BY total DESC
LIMIT 1;
```

billing_city	total
Prague	273.24

Insights –
a) Prague is the undisputed leader by a significant margin. With a total invoice sum of

273.24, it far surpasses the second-place city, Mountain View (169.29), by over 100. This isn't a close race; it's a decisive victory for Prague.

b) After Prague, there's a cluster of well-known tech and cultural hubs in the next tier: Mountain View, London, Berlin, and Paris. These all have totals above 150, indicating strong customer bases.

c) The top 10 includes cities from Europe (Prague, London, Berlin, Paris, Dublin), North America (Mountain View), and South America (São Paulo, São José dos Campos, Brasília). This shows a global customer base, but with a very strong European presence.

d) Throwing a festival in Prague is a strategic decision. You are investing in the city where your highest-value customers are already located. This rewards their loyalty and is very likely to generate significant additional revenue and brand engagement from your best market.

Q5. Who is the best customer ? The customer who has spent the most money will be declared the best customer. Write a query that returns the person who has spent the most money.

SELECT c.customer_id,
        c.first_name,
        c.last_name,
        ROUND(SUM(i.total),2) AS total
FROM customer c
JOIN invoice i ON c.customer_id = i.customer_id
GROUP BY c.customer_id,
          c.first_name,
          c.last_name
ORDER BY total DESC LIMIT 1;

customer_id	first_name	last_name	total
5	FrantiÅiek	WichterlovÃi	144.54

Insights –

a) František Wichterlová is clearly the best customer with a total spending of $144.54.

b) Top customers come from various regions including the Czech Republic (František), Ireland (Hugh), India (Manoj), Portugal (Luís), and Canada (François).

c) Spending Distribution:
- Only 3 customers spent over $140
- 10 customers spent between $90–$130
- Most customers spent between $60–$90
- 5 customers spent less than $50

Q6. Write a query to return the email, first name, last name, & Genre of all Rock Music listeners. Return your list ordered alphabetically by email starting with A.

SELECT DISTINCT email,
                first_name,
                last_name
FROM customer c
JOIN invoice i ON c.customer_id = i.customer_id
JOIN invoice_line il ON i.invoice_id = il.invoice_id
WHERE track_id IN (
                SELECT track_id
                FROM track t
                JOIN genre g ON g.genre_id = t.genre_id
                WHERE g.name LIKE 'Rock')
ORDER BY email;

email	first_name	last_name
aaronmitchell@yahoo.ca	Aaron	Mitchell
alero@uol.com.br	Alexandre	Rocha
astrid.gruber@apple.at	Astrid	Gruber
bjorn.hansen@yahoo.no	BjÃ¸rn	Hansen
camille.bernard@yahoo.fr	Camille	Bernard
daan_peeters@apple.be	Daan	Peeters
diego.gutierrez@yahoo.ar	Diego	GutiÃ©rrez
dmiller@comcast.com	Dan	Miller
dominiquelefebvre@gmail.com	Dominique	Lefebvre
edfrancis@yachoo.ca	Edward	Francis
eduardo@woodstock.com.br	Eduardo	Martins
ellie.sullivan@shaw.ca	Ellie	Sullivan

Q7. Let's invite the artists who have written the most rock music in our dataset. Write a query that returns the Artist name and total track.

```sql
SELECT ar.artist_id,
       ar.name,
       COUNT(t.track_id) AS num_of_songs
FROM artist ar
JOIN album al ON al.artist_id = ar.artist_id
JOIN track t ON t.album_id = al.album_id
WHERE t.genre_id IN (
    SELECT genre_id
    FROM genre
    WHERE name LIKE 'Rock')
GROUP BY ar.artist_id,
         ar.name
ORDER BY num_of_songs DESC;
```

artist_id	name	num_of_songs
1	AC/DC	18
3	Aerosmith	15
8	Audioslave	14
22	Led Zeppelin	14
4	Alanis Morissette	13
5	Alice In Chains	12
23	Frank Zappa & Captain Beefheart	9
2	Accept	4

Insights –
   a) AC/DC is the dominant rock artist in the dataset with 18 tracks, establishing them as a primary headline act candidate.
   b) Aerosmith, Audioslave, and Led Zeppelin form a strong second tier (14–15 tracks), making them ideal for other headline or major supporting slots.
   c) Artists like Alanis Morissette and Alice In Chains (12–13 tracks) have a substantial catalog, perfect for high-profile supporting acts to ensure a deep and engaging performance.
   d) The data shows a steep decline after the top 5-6 artists, indicating a clear divide between the most prolific rock acts and the rest.

Q8. Return all the track names that have a song length longer than the average song length. Return the Name and Milliseconds for each track. Order by the song length with the longest songs listed first.

SELECT name,
        milliseconds
FROM track
WHERE milliseconds > (
                    SELECT AVG(milliseconds) as avg_track_length
                    FROM track
)
ORDER BY milliseconds DESC;

name	milliseconds
How Many More Times	711836
Advance Romance	677694
Sleeping Village	644571
You Shook Me(2)	619467
Talkin' 'Bout Women Obviously	589531
Stratus	582086
No More Tears	555075
The Alchemist	509413
Wheels Of Confusion / The Straightener	494524
Book Of Thel	494393
You Oughta Know (Alternate)	491885
Terra	482429

Insights –
a) A significant portion of the listed songs far exceed the typical 3-4 minute pop song length, with many tracks exceeding 6.5 minutes and the longest ("How Many More Times") nearing 12 minutes.
b) The list is heavily populated by classic hard rock, heavy metal (e.g., Black Sabbath, Metallica), and progressive/jam bands (e.g., Frank Zappa, Led Zeppelin), genres known for extended solos and complex structures.
c) The presence of tracks like "You Shook Me (2)" and "The Beginning...At Last" suggests these may be alternate takes, live versions, or extended jams, which are inherently longer.
d) The prevalence of such long songs indicates the data is likely from album-focused collections where artistic expression is valued over commercial radio-friendly durations.
e) The average song length here is approximately 392,254 ms (~6.5 min), providing a clear benchmark. All returned tracks are longer than this, with the top results being substantially so.

Q9. Find how much amount is spent by each customer on artists. Write a query to return the customer name, artist name, and total spent.

```sql
WITH best_selling_artist AS (
    SELECT ar.artist_id,
            ar.name AS artist_name,
            SUM(il.unit_price * il.quantity) AS total_spent
    FROM invoice_line il
    JOIN track t ON t.track_id = il.track_id
    JOIN album al ON al.album_id = t.album_id
    JOIN artist ar ON ar.artist_id = al.artist_id
    GROUP BY 1,2
    ORDER BY 3 DESC)
SELECT c.customer_id,
        CONCAT(c.first_name," ",c.last_name) AS name,
        bsa.artist_name AS artist_name,
        ROUND(SUM(il.unit_price * il.quantity),2) AS total_spent
FROM invoice i
JOIN customer c ON c.customer_id = i.customer_id
JOIN invoice_line il ON il.invoice_id = i.invoice_id
JOIN track t ON t.track_id = il.track_id
JOIN album al ON al.album_id = t.album_id
JOIN best_selling_artist bsa ON bsa.artist_id = al.artist_id
GROUP BY 1, 2, 3
ORDER BY 4 DESC;
```

customer_id	name	artist_name	total_spent
54	Steve Murray	AC/DC	17.82
15	Jennifer Peterson	Aerosmith	14.85
55	Mark Taylor	Aerosmith	14.85
13	Fernanda Ramos	AntÃ´nio Carlos Jobim	13.86
2	Leonie KÃ¶hler	Audioslave	13.86
30	Edward Francis	Alanis Morissette	12.87
52	Emma Jones	Alanis Morissette	12.87
34	JoÃ£o Fernandes	Alanis Morissette	12.87
25	Victor Stevens	Alice In Chains	11.88
53	Phil Hughes	AC/DC	10.89
21	Kathy Chase	AC/DC	10.89
49	StanisÅ‚aw WÃ³jcik	Buddy Guy	10.89

Insights –

a) Customers like Steve Murray (AC/DC: $17.82), Jennifer Peterson, and Mark Taylor (Aerosmith: $14.85 each) are the highest individual spenders on a single artist.

b) AC/DC and Aerosmith are the top revenue-generating artists, with numerous high-value transactions from a large customer base.

c) Most customers spend on multiple artists. For example, Stanisław Wójcik spent on Buddy Guy ($10.89), AC/DC ($9.90), and others, indicating varied musical tastes.

d) While a few transactions are large (over $10), the vast majority are small purchases ($0.99), forming a long tail of revenue from many low-value sales across many artists.

Q10. We want to find out the most popular music Genre for each country. We determine the most popular genre as the genre with the highest amount of purchases. Write a query that returns each country along with the top Genre. For countries where the maximum number of purchases is shared, return all Genres.

```sql
WITH popular_genre AS (
  SELECT COUNT(il.quantity) AS purchases,
         c.country,
         g.name AS genre_name,
         ROW_NUMBER()
         OVER(PARTITION BY c.country ORDER BY COUNT(il.quantity) DESC) AS row_num
  FROM invoice_line il
  JOIN invoice i ON i.invoice_id = il.invoice_id
  JOIN customer c ON c.customer_id = i.customer_id
  JOIN track t ON t.track_id = il.track_id
  JOIN genre g ON g.genre_id = t.genre_id
  GROUP BY 2,3
  ORDER BY 2 ASC, 1 DESC)
SELECT country,
       genre_name,
       purchases
FROM popular_genre
WHERE row_num <= 1;
```

country	genre_name	purchases
Argentina	Rock	1
Australia	Rock	18
Austria	Rock	6
Belgium	Rock	5
Brazil	Rock	26
Canada	Rock	57
Chile	Rock	7
Czech Republic	Rock	14
Denmark	Rock	6
Finland	Rock	6
France	Rock	26
Germany	Rock	28
Hungary	Rock	4
India	Rock	13

Insights –

a) Rock is unequivocally the world's most popular music genre, being the top genre in 22 of the 24 countries analyzed.

b) The USA and Canada are the largest markets for rock music, with 70 and 57 purchases respectively, significantly outpacing other nations.

c) Major European economies like Germany, France, and the United Kingdom show strong rock music markets, further cementing its global appeal.

d) Only two countries break the rock trend: Spain's top genre is Metal (4 purchases), and Norway's is also Metal (2 purchases), indicating distinct regional musical preferences.

Q11. Write a query that determines the customer that has spent the most on music for each country. Write a query that returns the country along with the top customer and how much they spent. For countries where the top amount spent is shared, provide all customers who spent this amount.

```
WITH customer_with_country AS (
  SELECT  c.customer_id,
          CONCAT(first_name," ", last_name) AS name,
          billing_country,
          SUM(total) AS total_spent,
          ROW_NUMBER()
          OVER(PARTITION BY billing_country ORDER BY SUM(total) DESC) AS row_num
  FROM invoice i
JOIN customer c ON c.customer_id = i.customer_id
```

```
    GROUP BY 1,2,3
    ORDER BY 3, 4 DESC
)
SELECT customer_id,
        name,
        billing_country,
        ROUND(total_spent,2) AS total_spent
FROM customer_with_country
WHERE row_num = 1;
```

customer_id	name	billing_country	total_spent
56	Diego GutiÃ©rrez	Argentina	39.6
55	Mark Taylor	Australia	81.18
7	Astrid Gruber	Austria	69.3
8	Daan Peeters	Belgium	60.39
1	LuÃs GonÃ§alves	Brazil	108.9
3	FranÃ§ois Tremblay	Canada	99.99
57	Luis Rojas	Chile	97.02
5	FrantiÅ¡ek WichterlovÃ¡	Czech Republic	144.54
9	Kara Nielsen	Denmark	37.62
44	Terhi HÃ¤mÃ¤lÃ¤inen	Finland	79.2
42	Wyatt Girard	France	99.99
37	Fynn Zimmermann	Germany	94.05
45	Ladislav KovÃ¡cs	Hungary	78.21
58	Manoj Pareek	India	111.87

Insights –
a) Luis Rojas from Chile and Hugh O'Reilly from Ireland are the highest individual spenders globally, with totals of $97.02 and $114.84, respectively.
b) Three customers from different countries—Spain (Enrique Muñoz), the UK (Phil Hughes), and the USA (Jack Smith)—are tied for the top spot in their respective countries, each having spent exactly $98.01. Your query correctly included all three.
c) The highest spending amount varies significantly between countries, from a high of over $144 in the Czech Republic (František Wichterlová) to a low of under $38 in Denmark (Kara Nielsen), indicating vastly different market engagement levels.

Q12. List top 10 most popular artists ?

```
SELECT ar.name AS artist_name,
        COUNT(il.quantity) AS purchases
```

```
FROM invoice_line il
JOIN track t ON t.track_id = il.track_id
JOIN album al ON al.album_id = t.album_id
JOIN artist ar ON ar.artist_id = al.artist_id
GROUP BY 1
ORDER BY 2 DESC
LIMIT 10;
```

artist_name	purchases
AC/DC	124
Aerosmith	80
Alanis Morissette	75
Black Sabbath	68
Alice In Chains	59
Audioslave	43
Apocalyptica	37
Buddy Guy	23
AntÃ´nio Carlos Jobim	18
Black Label Society	15

Insights –

a) The list is overwhelmingly dominated by legendary rock and metal bands, with AC/DC, Aerosmith, and Led Zeppelin occupying the top spots, indicating a strong user preference for this genre.

b) AC/DC is the undisputed leader with a significant lead (124), far surpassing Aerosmith (80) in second place, suggesting a massive gap in listenership or sales data between the top two.

c) Despite rock's dominance, the inclusion of Brazilian legends like Antônio Carlos Jobim and Caetano Veloso shows an appreciable, though niche, audience for Bossa Nova and MPB (Brazilian Popular Music).

Q13. List top 10 most popular songs ?

```
SELECT t.name AS song_name,
    COUNT(il.quantity) AS purchases
FROM invoice_line il
JOIN track t ON t.track_id = il.track_id
GROUP BY 1
```

```
ORDER BY 2 DESC
LIMIT 10;
```

song_name	purchases
Put The Finger On You	13
Night Of The Long Knives	10
Snowballed	9
Mary Jane	9
Behind The Wall Of Sleep	9
Evil Woman	8
For Those About To Rock (We Salute You)	8
Dog Eat Dog	8
Deuces Are Wild	8
Sunshine	8

Insights –

a) "Put The Finger On You" is the undisputed number one, with a significant lead (13 plays) over the next contenders.

b) The next four songs ("Night Of The Long Knives," "Snowballed," "Mary Jane," "Behind The Wall Of Sleep") form a tight group with 9-10 plays, indicating a strong, consistent preference.

c) The list is characterized by frequent ties, with five songs sharing the 8-play count and four sharing the 7-play count, suggesting a diverse set of fan favorites beyond the top spot.

d) The popularity is measured by a simple play count metric, and the top songs show a noticeable drop from the peak of 13 down to 8, highlighting the top track's outlier status.

Q14. What are the average prices of different types of music ?

```
WITH purchases AS
    (SELECT g.name AS genre,
            SUM(total) AS total_spent
    FROM invoice i
    JOIN invoice_line il ON i.invoice_id = il.invoice_id
    JOIN track t ON t.track_id = il.track_id
    JOIN genre g ON g.genre_id = t.genre_id
    GROUP BY 1
    ORDER BY 2)
```

```sql
SELECT genre,
    CONCAT('$ ', ROUND(AVG(total_spent),2)) AS total_spent
FROM purchases
GROUP BY genre;
```

genre	total_spent
Reggae	$ 18.81
Alternative & Punk	$ 99
Latin	$ 144.54
Blues	$ 224.73
Jazz	$ 256.41
Metal	$ 940.5
Rock	$ 3664.98

Insights –

a) There is a massive range in average album prices across genres, from Reggae at $18.81 to Rock at $3,664.98. This suggests genre is a major pricing factor.

b) Rock music has the highest average price by a huge margin ($3,664.98), indicating it is likely the most profitable genre for the store or has a catalog of premium, high-value albums

c) Genres like Jazz ($256.41), Blues ($224.73), and Latin ($144.54) have moderately low averages, while Reggae ($18.81) and Punk/Alternative ($99) are at the bottom. This may reflect smaller audience demand or different pricing strategies.

d) The data reveals three tiers: a premium tier (Rock), a mid-range tier (Metal, Jazz, Blues), and a budget tier (Latin, Punk, Reggae), providing valuable insight for sales and marketing strategies.

Q15. What are the most popular countries for music purchases ?

```sql
SELECT c.country,
        COUNT(il.quantity) AS purchases
FROM invoice_line il
JOIN invoice i ON i.invoice_id = il.invoice_id
JOIN customer c ON c.customer_id = i.customer_id
GROUP BY country
ORDER BY purchases DESC;
```

country	purchases
USA	1051
Canada	541
Brazil	432
France	393
Germany	338
Czech Republic	276
United Kingdom	248
Portugal	187
India	185
Ireland	116
Spain	99
Chile	98
Australia	82

Insights –

a) The USA is the dominant market, with over 1,051 purchases, nearly double that of the next country, Canada (541). This establishes it as the primary driver of global music sales.

b) The top 10 list is heavily dominated by North American (USA, Canada) and European nations (France, Germany, Czech Republic, UK, Portugal, Ireland), indicating these regions have the highest customer engagement.

c) Significant markets emerge from South America, with Brazil (432) and Argentina (40) making the list, highlighting a strong regional presence.

d) The data reveals strong engagement from English-speaking countries (USA, Canada, UK, Ireland, Australia) and major European economies, suggesting a correlation between market size, language, and purchase volume.

e) The presence of the Czech Republic (276) and Portugal (187) in the top 10, ahead of larger economies like Australia, points to unexpectedly high per-capita music consumption in these regions.

Conclusion —

- The USA is the most popular country for music purchases with 1051 purchases followed by Canada and Brazil with 541 and 432 purchases respectively.
- Rock music is the top selling genre with $3664.98 spent and Reggae is the lowest selling genre.
- AC/DC is the top selling artist.
- Prague has the highest sales volume.