Music Store Playlist Analysis: SQL

Goal

This beginner's SQL project focuses on analyzing a Music Store's playlist data to optimize business decisions and drive growth. The objective is to thoroughly examine the store's data, providing actionable insights to enhance decision-making processes.

Project Outcomes

Key Outcomes:

1. Identify the most senior employee based on job title.
2. Determine which countries have the highest number of invoices.
3. Discover the top three highest invoice totals.
4. Identify the city with the highest sum of invoice totals.
5. Find the best customer based on total spending.
6. Identify Rock Music listeners and gather their details.
7. Discover the top Rock bands based on track count.
8. List tracks longer than the average song length.
9. Analyze customer spending on top artists.
10. Determine the most popular music genre per country.
11. Identify the top-spending customer in each country.

Benefits to the Music Store:

- Improved targeting for marketing campaigns.
- Enhanced customer loyalty programs.
- Better inventory and sales management.
- Optimized event planning and promotional activities.
- Insightful customer segmentation and profiling.

Query 1: Who is the senior most employee based on the job title?

Query:

```
SELECT employee_id,first_name, last_name, levels FROM employee
ORDER BY levels DESC LIMIT 1;
```

Output:

	employee_id [PK] character varying (50)	first_name character		last_name character	levels character varying (10)
1	9	Mohan	...	Madan	L7

The senior most customer of our Music Store is Mr. Madhav Mohan.

Benefit: Identifying the senior most employee helps in understanding leadership structure and decision-making hierarchy.

Query 2: Which countries have the most Invoices?

Query:

```
SELECT COUNT(*) as c, billing_country
FROM invoice
GROUP BY billing_country
ORDER BY c DESC;
```

Output:

	c bigint	billing_country character varying (30)
1	131	USA
2	76	Canada
3	61	Brazil
4	50	France
5	41	Germany
6	30	Czech Republic
7	29	Portugal
8	28	United Kingdom
9	21	India

USA, Canada and Brazil generate the most invoices with USA bagging the highest total count of 131.

Benefit: Understanding which countries generate the most invoices can guide international marketing and sales strategies.

Query 3: What are the top 3 values of total Invoices?

Query:

```
SELECT total FROM Invoice
ORDER BY total DESC
LIMIT 3;
```

Output:

	total double precision
1	23.759999999999998
2	19.8
3	19.8

Benefit: Identifying high-value transactions can help focus on high-spending customers and tailor premium services or products for them.

Query 4: Which city has the best customers? We would like to throw a promotional Music Festival in the city where we made the most money.

Query:

```
SELECT SUM(total) as invoice_total, billing_city
FROM Invoice
GROUP BY billing_city
ORDER BY invoice_total DESC;
```

Output:

	invoice_total double precision	billing_city character varying (30)
1	273.24000000000007	Prague
2	169.29	Mountain View
3	166.32	London
4	158.4	Berlin
5	151.47	Paris
6	129.69	São Paulo
7	114.83999999999997	Dublin
8	111.86999999999999	Delhi
9	108.89999999999998	São José dos Campos

The city with the highest invoice total is 'Prague' with the amount summing up to 273.

Benefit: Determining the city with the highest revenue can inform location-based promotions and events, such as a Music Festival.

Query 5: Who is the best customer? The customer who has spent the most money will be declared the best customer.

Query:

```
SELECT customer.customer_id, customer.first_name, customer.last_name,
SUM(invoice.total) as total
FROM Customer
JOIN invoice ON customer.customer_id = invoice.customer_id
GROUP BY customer.customer_id
ORDER BY total DESC
LIMIT 1;
```

Output:

	customer_id [PK] integer	first_name character	last_name character	total double precision
1	5	R ...	Madhav	144.54000000000002

The best customer turned out to be Mr. Madhav with his expenditure amount equaling 144.54.

Benefit: Recognizing the best customer can help in developing personalized engagement strategies to enhance loyalty and retention.

Query 6: Write query to return the email, first name, last name, & Genre of all Rock Music listeners.

Query:

```
SELECT DISTINCT email AS Email, first_name AS FirstName, last_name AS
LastName, genre.Name AS Name
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
JOIN invoice_line ON invoice_line.invoice_id = invoice.invoice_id
JOIN track ON track.track_id = invoice_line.track_id
JOIN genre ON genre.genre_id = track.genre_id
WHERE genre.name LIKE 'Rock'
ORDER BY email ASC;
```

Output:

	email character varying (50)	firstname character	lastname character	name character varying (120)
1	aaronmitchell@yahoo.ca	Aaron	Mitchell	Rock
2	alero@uol.com.br	Alexandre	Rocha	Rock
3	astrid.gruber@apple.at	Astrid	Gruber	Rock
4	bjorn.hansen@yahoo.no	Bjørn	Hansen	Rock
5	camille.bernard@yahoo.fr	Camille	Bernard	Rock
6	daan_peeters@apple.be	Daan	Peeters	Rock
7	diego.gutierrez@yahoo.ar	Diego	Gutiérrez	Rock
8	dmiller@comcast.com	Dan	Miller	Rock
9	dominiquelefebvre@gmail.c...	Dominique	Lefebvre	Rock

Benefit: Targeting Rock Music listeners for genre-specific marketing and events.

Query 7: Let's invite the artists who have written the most rock music in our dataset.

Query:

```
SELECT DISTINCT email, first_name, last_name
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
JOIN invoice_line ON invoice_line.invoice_id = invoice.invoice_id
WHERE track_id IN(
    SELECT track_id FROM track
    JOIN genre ON track.genre_id = genre.genre_id
    WHERE genre.name LIKE 'Rock'
)
ORDER BY email;
```

Output:

	email character varying (50)	first_name character	last_name character
1	aaronmitchell@yahoo.ca	Aaron	Mitchell
2	alero@uol.com.br	Alexandre	Rocha
3	astrid.gruber@apple.at	Astrid	Gruber
4	bjorn.hansen@yahoo.no	Bjørn	Hansen
5	camille.bernard@yahoo.fr	Camille	Bernard
6	daan_peeters@apple.be	Daan	Peeters
7	diego.gutierrez@yahoo.ar	Diego	Gutiérrez
8	dmiller@comcast.com	Dan	Miller
9	dominiquelefebvre@gmail.c...	Dominique	Lefebvre

Benefit: Inviting top Rock bands can enhance music events and attract genre-specific audiences.

Query 8: Return all the track names that have a song length longer than the average song length.

Query:

```
SELECT name, milliseconds
FROM track
WHERE milliseconds > (
    SELECT AVG(milliseconds) AS Avg_track_length
    FROM track
)
ORDER BY milliseconds DESC;
```

Output:

	name character varying (150)	milliseconds integer
1	Occupation / Precipice	5286953
2	Through a Looking Glass	5088838
3	Greetings from Earth, Pt. 1	2960293
4	The Man With Nine Lives	2956998
5	Battlestar Galactica, Pt. 2	2956081
6	Battlestar Galactica, Pt. 1	2952702
7	Murder On the Rising Star	2935894
8	Battlestar Galactica, Pt. 3	2927802
9	Take the Celestra	2927677

Benefit: Curating playlists with longer tracks can cater to specific customer preferences and enhance user experience.

Query 9: Find how much amount spent by each customer on artists? Write a query to return customer name, artist name and total spent.

Query:

```
WITH best_selling_artist AS (
    SELECT artist.artist_id AS artist_id, artist.name AS artist_Name,
    SUM(invoice_line.unit_price * invoice_line.quantity) AS total_sales
    FROM invoice_line
    JOIN track ON track.track_id = invoice_line.track_id
    JOIN album ON album.album_id = track.album_id
    JOIN artist ON artist.artist_id = album.artist_id
    GROUP BY 1
    ORDER BY 3 DESC
    LIMIT 1
)
SELECT c.customer_id, c.first_name, c.last_name, bsa.artist_name,
SUM(il.unit_price * il.quantity) AS amount_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 a ON a.album_id = t.album_id
JOIN best_selling_artist bsa ON bsa.artist_id = a.artist_id
GROUP BY 1, 2, 3, 4
ORDER BY 5 DESC;
```

Output:

	customer_id integer	first_name character	last_name character	artist_name character varying (120)	amount_spent double precision
1	56	Diego	Gutiérrez	Queen	16.830000000000002
2	8	Daan	Peeters	Queen	0.99
3	11	Alexandre	Rocha	Queen	0.99
4	17	Jack	Smith	Queen	0.99
5	48	Johannes	Van der Berg	Queen	0.99
6	51	Joakim	Johansson	Queen	0.99

The highest amount spent on an artist is $16.83, which was spent by Diego G. on Queen.

Benefit: Understanding customer spending on popular artists helps tailor recommendations and marketing efforts.

Query 10: 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 number of purchases.

Query:

```
WITH popular_genre AS
(
    SELECT COUNT(invoice_line.quantity) AS purchases, customer.country,
genre.name, genre.genre_id,
    ROW_NUMBER() OVER(PARTITION BY customer.country ORDER BY
COUNT(invoice_line.quantity) DESC) AS RowNo
    FROM invoice_line
    JOIN invoice ON invoice.invoice_id = invoice_line.invoice_id
    JOIN customer ON customer.customer_id = invoice.customer_id
    JOIN track ON track.track_id = invoice_line.track_id
    JOIN genre ON genre.genre_id = track.genre_id
    GROUP BY 2, 3, 4
    ORDER BY 2 ASC, 1 DESC
)
SELECT * FROM popular_genre WHERE RowNo <= 1;
```

Output:

	purchases bigint	country character varying (50)	name character varying (120)	genre_id character varying (50)	rowno bigint
1	17	Argentina	Alternative & Punk	4	1
2	34	Australia	Rock	1	1
3	40	Austria	Rock	1	1
4	26	Belgium	Rock	1	1
5	205	Brazil	Rock	1	1
6	333	Canada	Rock	1	1
7	61	Chile	Rock	1	1
8	143	Czech Republic	Rock	1	1
9	24	Denmark	Rock	1	1

Benefit: Identifying the most popular genres per country allows for culturally targeted promotions and inventory management.

Query 11: Write a query that determines the customer that has spent the most on music for each country.

Query:

```
WITH Customer_with_country AS (
```

```
    SELECT customer.customer_id, first_name, last_name, billing_country,
SUM(total) AS total_spending,
    ROW_NUMBER() OVER(PARTITION BY billing_country ORDER BY SUM(total) DESC)
AS RowNo
    FROM invoice
    JOIN customer ON customer.customer_id = invoice.customer_id
    GROUP BY 1, 2, 3, 4
    ORDER BY 4 ASC, 5 DESC
)
SELECT * FROM Customer_with_country WHERE RowNo <= 1;
```

Output:

	customer_id integer	first_name character	last_name character	billing_country character varying (30)	total_spending double precision	rowno bigint
1	56	Diego	Gutiérrez	Argentina	39.6	1
2	55	Mark	Taylor	Australia	81.18	1
3	7	Astrid	Gruber	Austria	69.3	1
4	8	Daan	Peeters	Belgium	60.38999999999999	1
5	1	Luís	Gonçalves	Brazil	108.89999999999998	1
6	3	François	Tremblay	Canada	99.99	1
7	57	Luis	Rojas	Chile	97.02000000000001	1
8	5	R	Madhav	Czech Republic	144.54000000000002	1
9	9	Kara	Nielsen	Denmark	37.61999999999999	1

Benefit: Identifying top-spending customers in each country aids in creating personalized engagement strategies and improving customer service.

Generated Insights

- **Targeted Marketing Campaigns**: Focus promotional efforts in the USA, Canada, and Brazil, where invoice generation is highest. Leverage the popularity of Rock music in these regions to tailor marketing messages, potentially increasing customer engagement by 15%.

- **Localized Promotions in Key Cities**: Develop city-specific promotions, particularly in Prague, which shows the highest invoice total, and engage local influencers to drive sales by up to 10%.

- **Customer Loyalty Programs**: Introduce a loyalty program targeting high-value customers like Mr. Madhav, who has the highest expenditure, to enhance retention and encourage repeat purchases, potentially boosting sales by 20%.

- **Inventory Optimization**: Prioritize stocking and promoting Rock genre music and top artists like Queen, ensuring availability aligns with customer preferences, which could reduce stockouts and increase sales by 25%.

- **Data-Driven Decision Making**: Utilize detailed customer preferences and behaviors to refine inventory and sales strategies, improving decision-making processes and driving growth by up to 30%.

Questions and Discussion

Thank You!

Any Questions?