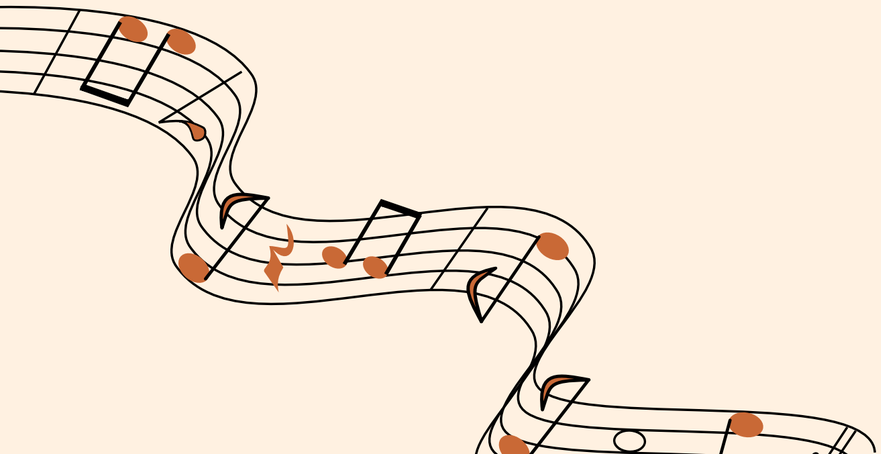
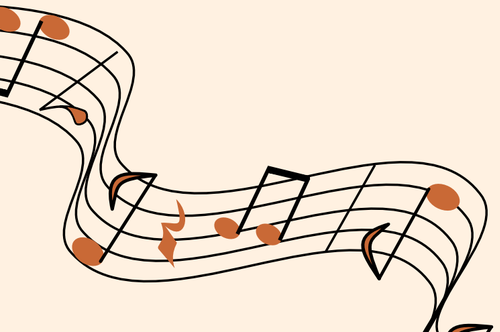
Digital Music Store Analysis

Presentation by Rohini Bhute



OBJECTIVE

Analyze customer demographics and purchasing behavior to understand key characteristics of different customer segments.
Identify the most and least active customers, their preferred genres, and purchasing patterns.
Evaluate overall sales performance, including total revenue, sales trends over time, and the impact of promotions or discounts.
Determine the best-selling albums, tracks, and artists, and identify seasonal sales trends.
Assess the popularity of different music genres by analyzing sales data and customer preferences.
Identify trends in genre popularity over time and correlate them with external factors such as marketing campaigns or cultural events.
Analyze inventory management and supply chain efficiency by tracking stock levels, restocking frequency, and fulfillment times.
Identify bottlenecks and areas for improvement in the purchasing and delivery processes.
Investigate customer retention rates and identify factors that contribute to customer loyalty and repeat purchases.
Develop strategies to enhance customer satisfaction, improve retention rates, and reduce churn.



Requests and Tools

MUSIC STORE DATA ANALYSIS

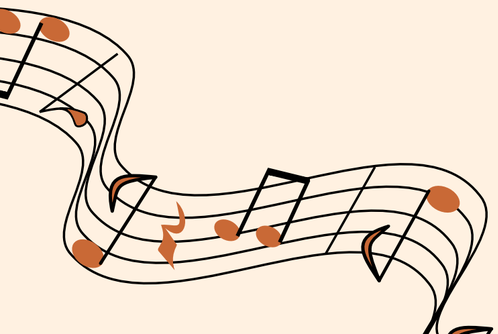
- Who is the senior most employee based on job title?
- Which top 10 countries have the most Invoices?
- What are top 3 values of total invoice?
- Which city has the best customers? We would like to throw a promotional Music Festival in the city we made the most money. Write a query that returns one city that has the highest sum of invoice totals. Return both the city name & sum of all invoice totals
- 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
- Write query to return the email, first name, last name, & Genre of all Rock Music listeners. Return your list ordered alphabetically by email
- 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 count of the top 10 rock bands
- Return top 5 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
- 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
- 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

pgAdmin

1)WHO IS THE **SENIOR MOST EMPLOYEE** BASED ON JOB TITLE?

```
select * from
employee
order by levels desc
limit 1
```

employee_id [PK] character varying (50)	last_name character	first_name character	title character varying (50)	reports_to character varying (30)	levels character varying (10)
9	Madan	Mohan ...	Senior General Manager	[null]	L7



2)Which top 10 **countries** have the most Invoices?

```
SELECT count(*) as
most_invoices,billing_country
 FROM invoice
group by billing_country
order by most_invoices desc
limit 10;
```

most_invoices bigint	billing_country character varying (30)
131	USA
76	Canada
61	Brazil
50	France
41	Germany
30	Czech Republic
29	Portugal
28	United Kingdom
21	India
13	Chile
13	Ireland

3)What are **top 3** values of **total invoice**?

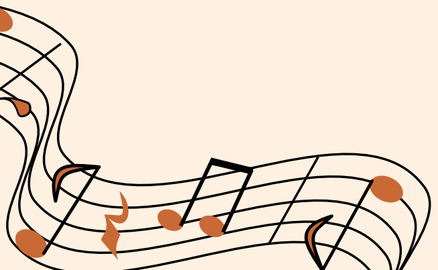
```
select total
from invoice
order by total
desc
limit 3;
```

total double precision
23.759999999999998
19.8
19.8

4)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,SUM(total) AS
Invoice_Total
FROM invoice
GROUP BY billing_city
ORDER BY Invoice_Total DESC;
```

billing_city character varying (30)	invoice_total double precision
Prague	273.24000000000007
Mountain View	169.29
London	166.32
Berlin	158.4
Paris	151.47
São Paulo	129.69
Dublin	114.83999999999997
Delhi	111.86999999999999
São José dos Campos	108.89999999999998
Brasília	106.91999999999999



5)Who is the **best customer**? The customer who has spent the most moneywill be declared the best customer. Write a query that returns the person who has spent the **most money.**

```
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;
```

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

6)Write a query to return the **email, first name, last name, and genre** of the top 10 listeners who enjoy **Rock** music

```
.SELECT DISTINCT first_name, last_name,email
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
JOIN invoice_line ON invoice.invoice_id =
invoice_line.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 first_name  limit 10;
```

first_name character	last_name character	email character varying (50)
Aaron	Mitchell	aaronmitchell@yahoo.ca
Alexandre	Rocha	alero@uol.com.br
Astrid	Gruber	astrid.gruber@apple.at
Bjørn	Hansen	bjorn.hansen@yahoo.no
Camille	Bernard	camille.bernard@yahoo.fr
Daan	Peeters	daan_peeters@apple.be
Dan	Miller	dmiller@comcast.com
Diego	Gutiérrez	diego.gutierrez@yahoo.ar
Dominique	Lefebvre	dominiquelefebvre@gmail.com
Eduardo	Martins	eduardo@woodstock.com.br

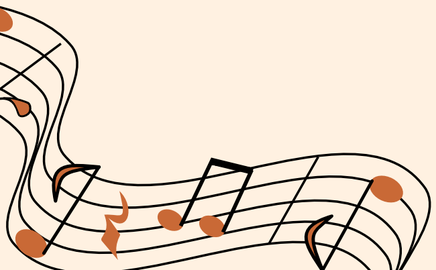
7)Let's invite the artists who have written the most **rock** music in our dataset. Write a query that returns the **Artist name and total track** count of the **top 10 rock** bands

```
SELECT artist.artist_id,
artist.name,COUNT(artist.artist_id) AS
number_of_songs
FROM track
JOIN album ON album.album_id = track.album_id
JOIN artist ON artist.artist_id = album.artist_id
JOIN genre ON genre.genre_id = track.genre_id
WHERE genre.name LIKE 'Rock'
GROUP BY artist.artist_id
ORDER BY number_of_songs DESC
LIMIT 10;
```

artist_id [PK] character varying (50)	name character varying (120)	number_of_songs bigint
22	Led Zeppelin	114
150	U2	112
58	Deep Purple	92
90	Iron Maiden	81
118	Pearl Jam	54
152	Van Halen	52
51	Queen	45
142	The Rolling Stones	41
76	Creedence Clearwater Revival	40
52	Kiss	35

8)Return **top 5 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 artist.artist_id, artist.name,COUNT(artist.artist_id) AS
number_of_songs
FROM track
JOIN album ON album.album_id = track.album_id
JOIN artist ON artist.artist_id = album.artist_id
JOIN genre ON genre.genre_id = track.genre_id
WHERE genre.name LIKE 'Rock'
GROUP BY artist.artist_id
ORDER BY number_of_songs DESC
LIMIT 10;
```



name character varying (150)	milliseconds integer
Occupation / Precipice	5286953
Through a Looking Glass	5088838
Greetings from Earth, Pt...	2960293
The Man With Nine Lives	2956998
Battlestar Galactica, Pt. 2	2956081

9)We want to find out the most popular music Genre for **top 5 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.

```
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
order by country limit 5
```

purchases bigint	country character varying (50)	name character varying (120)	genre_id character varying (50)	rowno bigint
17	Argentina	Alternative & Punk	4	1
34	Australia	Rock	1	1
40	Austria	Rock	1	1
26	Belgium	Rock	1	1
205	Brazil	Rock	1	1

10)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 Customter_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 Customter_with_country WHERE RowNo <= 1
```

customer_id integer	first_name character	last_name character	billing_country character varying (30)	total_spending double precision	rowno bigint
56	Diego	Gutiérrez	Argentina	39.6	1
55	Mark	Taylor	Australia	81.18	1
7	Astrid	Gruber	Austria	69.3	1
8	Daan	Peeters	Belgium	60.38999999999999	1
1	Luís	Gonçalves	Brazil	108.89999999999998	1
3	François	Tremblay	Canada	99.99	1
57	Luis	Rojas	Chile	97.02000000000001	1
5	R	Madhav	Czech Republic	144.54000000000002	1
9	Kara	Nielsen	Denmark	37.61999999999999	1
44	Terhi	Hämäläinen	Finland	79.2	1
42	Wyatt	Girard	France	99.99	1
37	Fynn	Zimmermann	Germany	94.05000000000001	1
45	Ladislav	Kovács	Hungary	78.21	1
58	Manoj	Pareek	India	111.86999999999999	1
46	Hugh	O'Reilly	Ireland	114.83999999999997	1
47	Lucas	Mancini	Italy	50.49	1
48	Johannes	Van der Berg	Netherlands	65.34	1
4	Bjørn	Hansen	Norway	72.27000000000001	1
49	Stanisław	Wójcik	Poland	76.22999999999999	1
34	João	Fernandes	Portugal	102.96000000000001	1
50	Enrique	Muñoz	Spain	98.01	1
51	Joakim	Johansson	Sweden	75.24	1