
MySQL Project On Music Store Data Analysis



OBJECTIVE

This project helps us to analyze the music playlist database. You can easily examine the dataset with SQL and help the store to understand its business growth by answering simple Questions.



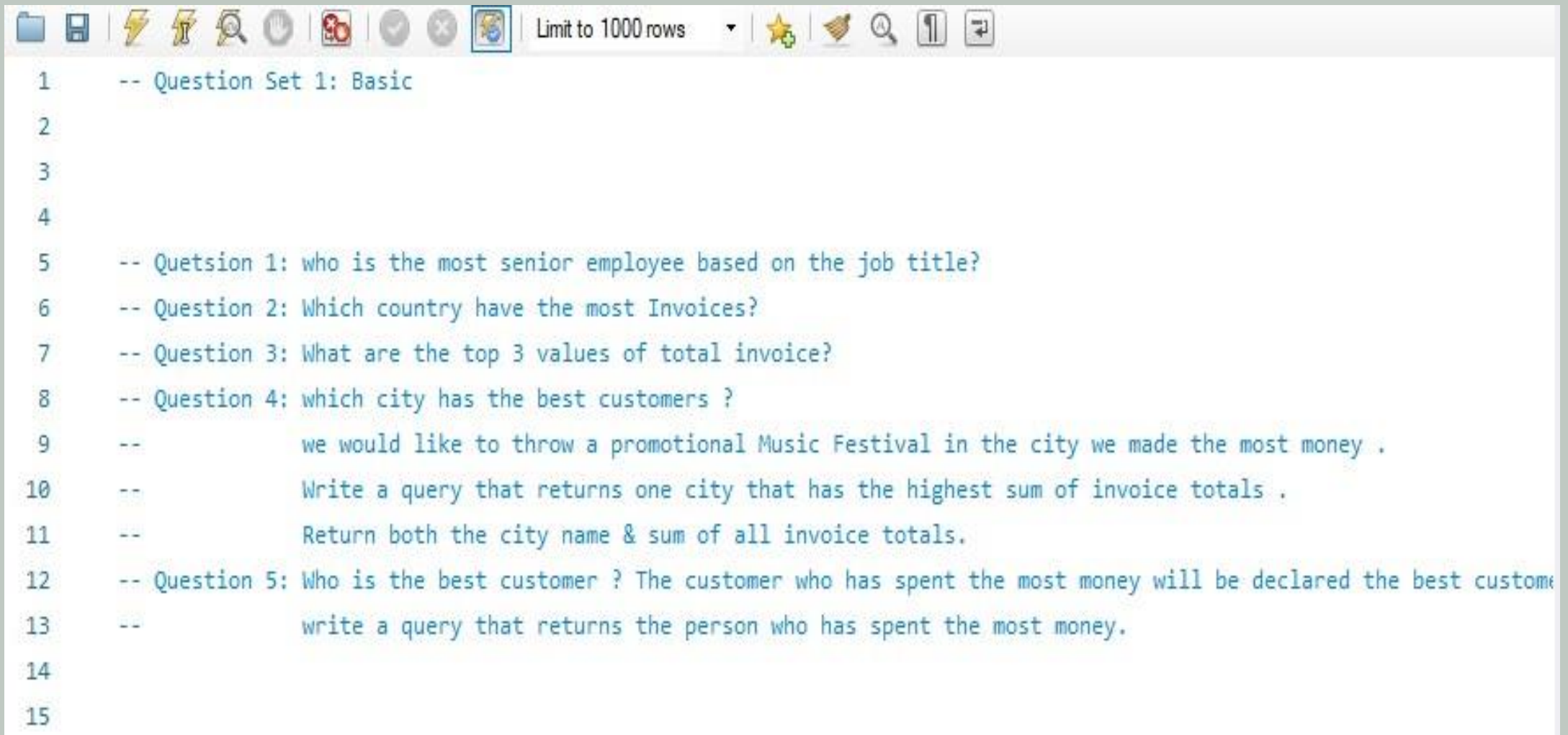
Introduction

Hello!

My name is Mansi, and in this project I have utilized MySQL database for analyzing the dataset of a music store provided to me .

Sql is a programming language, that allows the users to store, manipulate, and retrieve data efficiently in relational database





```
1  -- Question Set 1: Basic
2
3
4
5  -- Quetsion 1: who is the most senior employee based on the job title?
6  -- Question 2: Which country have the most Invoices?
7  -- Question 3: What are the top 3 values of total invoice?
8  -- Question 4: which city has the best customers ?
9  --           we would like to throw a promotional Music Festival in the city we made the most money .
10 --           Write a query that returns one city that has the highest sum of invoice totals .
11 --           Return both the city name & sum of all invoice totals.
12 -- Question 5: Who is the best customer ? The customer who has spent the most money will be declared the best customer
13 --           write a query that returns the person who has spent the most money.
14
15
```



```
1  -- Question 2: Which country have the most Invoices?
2
3  -- select * from music_database.invoice
4  •  SELECT
5      COUNT(billing_country) AS count, billing_country
6  FROM
7      invoice
8  GROUP BY billing_country
9  ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	count	billing_country
▶	131	USA
	76	Canada
	61	Brazil
	50	France
	41	Germany
	30	Czech Republic
	29	Portugal
	28	United Kingdom

Result 1 x ! Read Only



```
1  -- Question set 2 - Moderate
2
3  -- Question 1 : Write a query to return the email , first name , last name & genre  of all rock music listeners.
4  --           Return your list ordered alphabetically by email starting with A
5  -- Question 2 : Let's invite the artists who have written the most rock music in our dataset .
6  --           Write a query that returns the artist name and total track count of the top 10 rock bands
7  -- Question 3 : Return all the track names that have a song length  longer than the average song length .
8  --           Return the name and milliseconds for each track . order by the song length with the longest songs
9  --           listed first.
```

```

1  -- Question 1 : Write a query to return the email , first name , last name & genre of all rock music listeners.
2  --          Return your list ordered alphabetically by email starting with A
3  • SELECT DISTINCT
4      email, first_name, last_name
5  FROM
6      customer
7      JOIN
8      invoice ON customer.customer_id = invoice.customer_id
9      JOIN
10     invoice_line ON invoice.invoice_id = invoice_line.invoice_id
11 WHERE
12     track_id IN (SELECT
13         track_id
14         FROM
15             track
16             JOIN
17             genre ON track.genre_id = genre.genre_id
18             WHERE
19                 genre.name LIKE 'rock')
20 ORDER BY email;
21
22
23

```

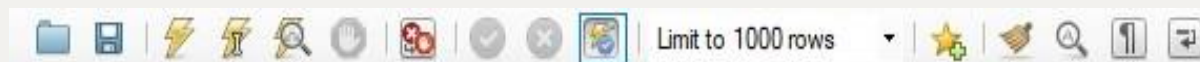
Result Grid

Filter Rows:

Export:

	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	Björn	Hansen
	camille.bernard@yahoo.fr	Camille	Bernard
	daan.peeters@apple.be	Daan	Peeters
	diego.gutierrez@yahoo.ar	Diego	Gutiérrez
	dmiller@comcast.com	Dan	Miller
	dominiquelefebvre@gmail.com	Dominique	Lefebvre
	edfrancis@yahoo.ca	Edward	Francis
	eduardo@woodstock.com.br	Eduardo	Martins
	ellie.sullivan@shaw.ca	Ellie	Sullivan
	emma.jones@hotmail.com	Emma	Jones
	enrique_munoz@yahoo.es	Enrique	Muñoz
	fernadaramos4@uol.com.br	Fernanda	Ramos
	fharris@google.com	Frank	Harris
	fralston@gmail.com	Frank	Ralston
	frantisekw@jetbrains.com	František	Wichterlov
	ftremblay@gmail.com	François	Tremblay

Result 2 x



```
1  -- Question 2 : Let's invite the artists who have written the most rock music in our dataset .
2  --           Write a query that returns the artist name and total track count of the top 10 rock bands
3
4  • SELECT
5      artist.artist_id,
6      artist.name,
7      COUNT(artist.artist_id) AS number_of_songs
8  FROM
9      track
10     JOIN
11     album2 ON album2.album_id = track.album_id
12     JOIN
13     artist ON artist.artist_id = album2.artist_id
14     JOIN
15     genre ON genre.genre_id = track.genre_id
16 WHERE
17     genre.name LIKE 'Rock'
18 GROUP BY artist.artist_id , artist.name
19 ORDER BY number_of_songs DESC
20 LIMIT 10;
```

	artist_id	name	number_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


```
1  -- Question 3 : Return all the track names that have a song length longer than the average song length . Return
2  -- the name and milliseconds for each track.Order by the song length with the longest songs listed first.
3
4  ●  SELECT
5      name, milliseconds
6  FROM
7      track
8  WHERE
9      milliseconds > (SELECT
10         AVG(milliseconds) AS avg_track_length
11         FROM
12             track)
13  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

```
1  -- Question Set 3- Advanced
2
3  -- Question 1 : Find How much amount spent by each customer on artists? write a query to return
4  --              return customer name , artist name,  and total spent.
5  -- Question 2 : We want to find out the most popular music Genre for each country .
6  --              We determine the most popular genre as the genre with the highest amount of purchases .
7  --              Write a query that returns each country along with the top Genre .
8  --              For countries where the maximum number of purchases is shared return all genres.
9  -- Question 3 : Write the query that determines the customer that has spent the most on music for each country .
10 --              write a query that returns the country along with the top customer and how much they spent .
11 --              For countries where the top amount spent is shared , provide all customers who spend this amount.
12
```

Limit to 1000 rows

```

1  -- Question 1 : Find How much amount spent by each customer on artists? write a query to return, customer name,
2  --               artist name, and total spent.
3
4  WITH best_selling_artist as (
5      SELECT  artist.artist_id as artist_id, artist.name as artist_name,
6              sum(invoice_line.unit_price*invoice_line.quantity) as total_sales
7      from invoice_line
8      join track on track.track_id = invoice_line.track_id
9      join album2 on album2.album_id = track.album_id
10     join artist on artist.artist_id = album2.artist_id
11     group by 1, 2
12     order by 3 desc
13     limit 1
14 )
15 Select c.customer_id, c.first_name, c.last_name, bsa.artist_name,
16        sum( il.unit_price*il.quantity) as amount_spent
17 from invoice i
18 join customer c on c.customer_id = i.customer_id
19 join invoice_line il on il.invoice_id = i.invoice_id
20 join track t on t.track_id = il.track_id
21 join album2 alb on alb.album_id = t.album_id
22 join best_selling_artist bsa on bsa.artist_id = alb.artist_id
23 group by 1,2,3,4
24 order by 5 desc;

```

Result Grid Filter Rows: Exports: Wrap Cell Contents:

	customer_id	first_name	last_name	artist_name	amount_spent
	54	Steve	Murray	AC/DC	17.82
	53	Phil	Hughes	AC/DC	10.89
	21	Kathy	Chase	AC/DC	10.89
	49	Stanisław	Wójcik	AC/DC	9.9
	1	Luís	Gonçalves	AC/DC	7.920000000000001
	24	Frank	Ralston	AC/DC	7.920000000000001
	31	Martha	Silk	AC/DC	3.96
	16	Frank	Harris	AC/DC	2.9699999999999998
	42	Wyatt	Girard	AC/DC	2.9699999999999998
	6	Helena	Holm ½	AC/DC	2.9699999999999998
	38	Niklas	Schröder	AC/DC	2.9699999999999998
	35	Madalena	Sampaio	AC/DC	2.9699999999999998
	44	Terhi	Hämäläinen	AC/DC	2.9699999999999998
	9	Kara	Nielsen	AC/DC	1.98
	34	João	Fernandes	AC/DC	1.98
	57	Luis	Rojas	AC/DC	1.98
	27	Patrick	Gray	AC/DC	1.98
	20	Dan	Miller	AC/DC	1.98
	30	Edward	Francis	AC/DC	1.98

Result 2

Limit to 1000 rows

```

1  -- Question 2 : We want to find out the most popular music Genre for each country. We determine the most popular
2  --               genre as the genre with the highest amount of purchases . Write a query that returns each country
3  --               along with the top Genre . For countries where the maximum number of purchases is shared return all
4  --               genres.
5
6  • with recursive
7  sales_per_country as (
8      select count(invoice_line.quantity) as purchases_per_genre, customer.country , genre.name, genre.genre_id
9      from invoice_line
10     join invoice on invoice.invoice_id = invoice_line.invoice_id
11     join customer on customer.customer_id = invoice.customer_id
12     join track on track.track_id = invoice_line.track_id
13     join genre on genre.genre_id = track.genre_id
14     group by 2, 3, 4
15     order by 2
16 ),
17 max_genre_per_country as (select max(purchases_per_genre) as max_genre_number, country
18     from sales_per_country
19     group by 2
20     order by 2)
21 SELECT sales_per_country.* FROM sales_per_country
22 JOIN max_genre_per_country ON sales_per_country.country = max_genre_per_country.country
23 WHERE sales_per_country.purchases_per_genre = max_genre_per_country.max_genre_number;
24

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	purchases_per_genre	country	name	genre_id
▶ 1		Argentina	Rock	1
18		Australia	Rock	1
6		Austria	Rock	1
5		Belgium	Rock	1
26		Brazil	Rock	1
57		Canada	Rock	1
7		Chile	Rock	1
14		Czech Republic	Rock	1
6		Denmark	Rock	1
6		Finland	Rock	1
26		France	Rock	1
28		Germany	Rock	1
4		Hungary	Rock	1
13		India	Rock	1
2		Ireland	Rock	1
3		Italy	Rock	1
6		Netherlands	Rock	1

Result 4 x


```

1  -- Question 3 : Write the query that determines the customer that has spent the most on music for each country .
2  --               write a query that returns the country along with the top customer and how much they spent .
3  --               For countries where the top amount spent is shared , provide all customers who spend this amount.
4
5  • with recursive
6  ⊖ customer_with_country as(
7      select customer.customer_id, customer.first_name, customer.last_name, billing_country,
8      sum(total) as total_spending
9      from invoice
10     join customer on customer.customer_id = invoice.customer_id
11     group by 1,2,3,4
12     order by 1,5 desc),
13
14  ⊖ country_max_spending as (
15      select billing_country, max(total_spending) as max_spending
16      from customer_with_country
17      group by billing_country)
18
19  SELECT cc.billing_country, cc.total_spending, cc.first_name, cc.last_name, cc.customer_id
20  FROM customer_with_country cc
21  JOIN country_max_spending ms ON cc.billing_country = ms.billing_country
22  WHERE cc.total_spending = ms.max_spending
23  ORDER BY 1;
24

```

Result Grid	Filter Rows:	Export:	Wrap Cell Contents:		
	billing_country	total_spending	first_name	last_name	customer_id
▶	Argentina	39.6	Diego	Gutiérrez	56
	Australia	81.18	Mark	Taylor	55
	Austria	69.3	Astrid	Gruber	7
	Belgium	60.38999999999999	Daan	Peeters	8
	Brazil	108.89999999999998	Luís	Gonçalves	1
	Canada	99.99	François	Tremblay	3
	Chile	97.02000000000001	Luis	Rojas	57
	Czech Republic	144.54000000000002	František	Wichterlová	5
	Denmark	37.61999999999999	Kara	Nielsen	9
	Finland	79.2	Terhi	Hämäläinen	44
	France	99.99	Wyatt	Girard	42
	Germany	94.05000000000001	Fynn	Zimmermann	37
	Hungary	78.21	Ladislav	Kovács	45
	India	111.86999999999999	Manoj	Pareek	58
	Ireland	114.83999999999997	Hugh	O'Reilly	46
	Italy	50.49	Lucas	Mancini	47
	Netherlands	65.34	Johannes	Van der Berg	48
	Norway	72.27000000000001	Bjørn	Hansen	4
	Poland	76.22999999999999	Stanisław	Wójcik	49

Result 2 x

Thank You