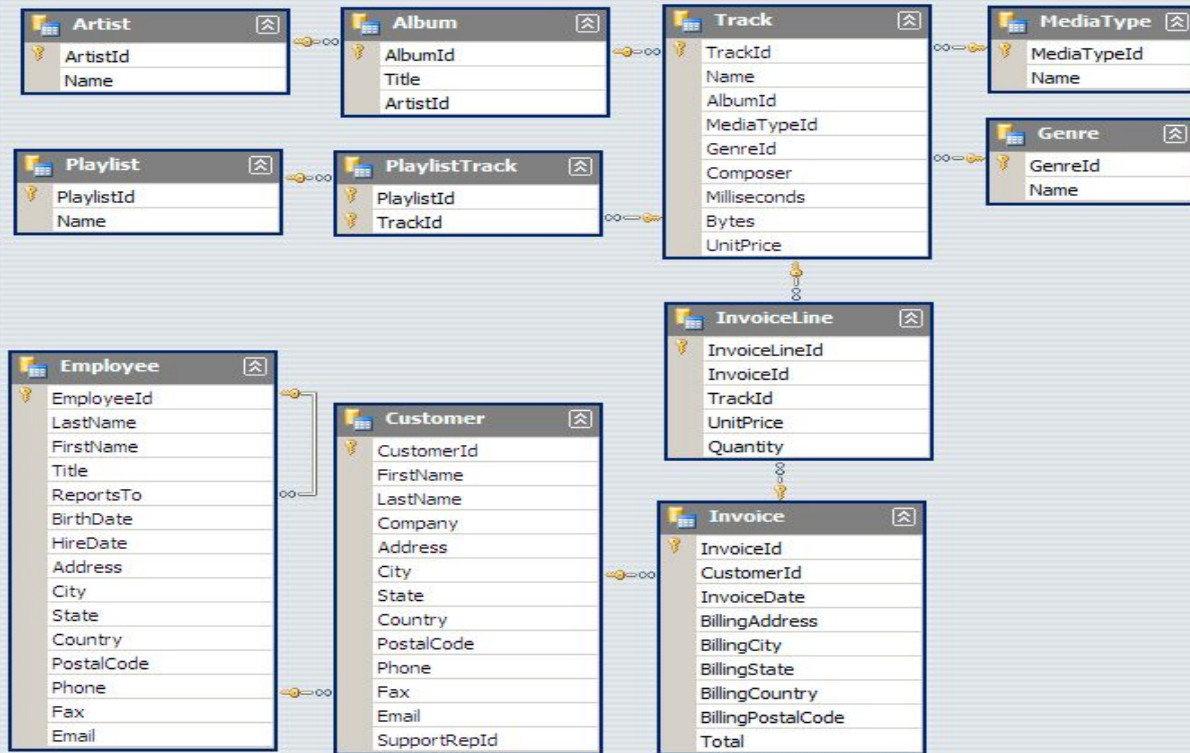




SQL DATA ANALYST PROJECT

By Md Rahmat Ali

Database schema



Questions Set 1

Write query to return the email, first name, last name, where Genre of users are Rock Music listeners. Return your list ordered alphabetically by email starting with A.

Query:

```
Query  Query History
1  Select distinct c.first_name, c.last_name, c.email
2  from customer as c
3  join invoice
4  on c.customer_id=invoice.customer_id
5  join invoice_line
6  on invoice.invoice_id = invoice_line.invoice_id
7  where invoice_line.track_id in(Select track.track_id
8  from track join genre
9  on track.genre_id = genre.genre_id
10 where genre.name = 'Rock')
11 order by c.email;
```

Output:

	first_name character		last_name character		email character varying (50)
1	Aaron	...	Mitchell	...	aaronmitchell@yahoo.ca
2	Alexandre	...	Rocha	...	alero@uol.com.br
3	Astrid		Gruber	...	astrid.gruber@apple.at
4	Bjorn		Hansen	...	bjorn.hansen@yahoo.no
5	Camille	...	Bernard	...	camille.bernard@yahoo.fr
6	Daan		Peeters	...	daan_peeters@apple.be
7	Diego		Gutiérrez	...	diego.gutierrez@yahoo.ar
8	Dan		Miller		dmiller@comcast.com
9	Dominique	...	Lefebvre	...	dominiquelefebvre@gmail.c...
10	Edward	...	Francis	...	edfrancis@yahoo.ca
11	Eduardo	...	Martins	...	eduardo@woodstock.com.br
Total rows: 59 of 59 Query complete 00:00:00.078					

Questions Set 1

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.

Query:

```
Query  Query History
1  WITH RockTracks AS (
2      SELECT track.album_id, album.artist_id
3      FROM track
4      JOIN album ON album.album_id = track.album_id
5      JOIN genre ON genre.genre_id = track.genre_id
6      WHERE genre.name = 'Rock'
7  )
8  SELECT artist.artist_id, artist.name, COUNT(*) AS n
9  FROM RockTracks
10 JOIN artist ON artist.artist_id = RockTracks.artist_id
11 GROUP BY artist.artist_id, artist.name
12 ORDER BY n DESC
13 LIMIT 10;
```

Output:

	artist_id [PK] character varying (50)	name character varying (120)	n bigint
1	22	Led Zeppelin	114
2	150	U2	112
3	58	Deep Purple	92
4	90	Iron Maiden	81
5	118	Pearl Jam	54
6	152	Van Halen	52
7	51	Queen	45
8	142	The Rolling Stones	41
9	76	Creedence Clearwater Revival	40
10	52	Kiss	35
Total rows: 10 of 10		Query complete 00:00:00.058	

Questions Set 1

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.

Query:

```
Query  Query History
1  ✓  Select name, milliseconds
2      from track
3      where milliseconds >
4          (Select avg(milliseconds) from track)
5      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
10	Fire In Space	2926593
Total rows: 494 of 494		Query complete 00:00

Questions Set 2

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

Query:

```
Query  Query History
1  WITH selling_artist as (
2      Select artist.artist_id, artist.name,
3      sum(invoice_line.unit_price*invoice_line.quantity)
4      from invoice_line
5      join track on track.track_id = invoice_line.track_id
6      join album on album.album_id = track.album_id
7      join artist on artist.artist_id = album.artist_id
8      group by artist.artist_id
9      order by 3 desc
10     limit 1
11 )
12 Select c.customer_id, c.first_name, c.last_name, selling_artist.name,
13 sum(il.unit_price*il.quantity)
14 from invoice as ins
15 join customer as c
16 on ins.customer_id = c.customer_id
17 join invoice_line il on il.invoice_id = ins.invoice_id
18 join track on track.track_id = il.track_id
19 join album on album.album_id = track.album_id
20 join selling_artist on selling_artist.artist_id = album.artist_id
21 Group by 1,2,3,4
22 order by 5 desc
```

Output:

	customer_id integer	first_name character	last_name character	name character var	sum double precis
1	46	Hugh ...	O'Reilly ...	Queen	27.71999999
2	38	Niklas ...	Schröder ...	Queen	18.81
3	3	François ...	Tremblay ...	Queen	17.82
4	34	João ...	Fernande...	Queen	16.83000000
5	53	Phil ...	Hughes ...	Queen	11.88
6	41	Marc ...	Dubois ...	Queen	11.88
7	47	Lucas ...	Mancini ...	Queen	10.89
8	33	Ellie ...	Sullivan ...	Queen	10.89
9	20	Dan ...	Miller ...	Queen	3.96
10	5	R ...	Madhav ...	Queen	3.96
Total rows: 43 of 43		Query complete 00:00:00.111			

Questions Set 2

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.

Query:

Output:

```
Query Query History
1  With famous_genre as (
2      Select i.billing_country, g.name, count(il.quantity) as purchases,
3      Row_Number() over(partition by i.billing_country
4      |order by count(il.quantity) desc) x
5      from invoice i
6      join customer on customer.customer_id = i.customer_id
7      join invoice_line il on i.invoice_id=il.invoice_id
8      join track t on il.track_id = t.track_id
9      join genre g on g.genre_id = t.genre_id
10     group by 1,2
11     order by 1 asc, 3 desc)
12
13 Select billing_country, name,
14 purchases from famous_genre where x<2
```

	billing_country character varying (30) 🔒	name character varying (120) 🔒	purchases bigint 🔒
1	Argentina	Alternative & Punk	17
2	Australia	Rock	34
3	Austria	Rock	40
4	Belgium	Rock	26
5	Brazil	Rock	205
6	Canada	Rock	333
7	Chile	Rock	61
8	Czech Republic	Rock	143
9	Denmark	Rock	24
10	Finland	Rock	46
Total rows: 24 of 24		Query complete 00:00:00.113	

Questions Set 2

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.

Query:

```
Query  Query History
1  WITH RECURSIVE
2      customter_with_country AS (
3          SELECT customer.customer_id,first_name,
4              last_name,billing_country,SUM(total) AS total_spending
5          FROM invoice
6          JOIN customer ON customer.customer_id = invoice.customer_id
7          GROUP BY 1,2,3,4
8          ORDER BY 2,3 DESC),
9
10     country_max_spending AS(
11         SELECT billing_country,MAX(total_spending) AS max_spending
12         FROM customter_with_country
13         GROUP BY billing_country)
14
15     SELECT cc.billing_country, cc.total_spending, cc.first_name,
16         cc.last_name, cc.customer_id
17     FROM customter_with_country cc
18     JOIN country_max_spending ms
19     ON cc.billing_country = ms.billing_country
20     WHERE cc.total_spending = ms.max_spending
21     ORDER BY 1;
```

Output:

	billing_country character varying (3)	total_spending double precision	first_name character	last_name character	customer_id integer
1	Argentina	39.6	Diego ...	Gutiérrez ...	56
2	Australia	81.18	Mark ...	Taylor ...	55
3	Austria	69.3	Astrid ...	Gruber ...	7
4	Belgium	60.38999999999999	Daan ...	Peeters ...	8
5	Brazil	108.89999999999998	Luís ...	Gonçalves ...	1
6	Canada	99.99	François ...	Tremblay ...	3
7	Chile	97.02000000000001	Luis ...	Rojas ...	57
8	Czech Republic	144.54000000000002	R ...	Madhav ...	5
9	Denmark	37.61999999999999	Kara ...	Nielsen ...	9
10	Finland	79.2	Terhi ...	Hämäläinen ...	44
Total rows: 24 of 24		Query complete 00:00:00.172			

Questions Set 3

Write a query to retrieve the senior most employee based on job title?

Query:

```
Query  Query History
1  ✓  SELECT title, last_name, first_name
2     FROM employee
3     ORDER BY levels DESC
4     LIMIT 1|
```

Output:

	title character varying (50) 🔒	last_name character 🔒	first_name character 🔒
1	Senior General Manager	Madan	Mohan ...

Questions Set 3

Which countries have the most Invoices and what are top 3 values of total invoice?

Query:

```
Query  Query History
1  SELECT billing_country, count(1) as c
2  from invoice
3  group by billing_country
4  order by c desc
5  limit 3
```

```
Query  Query History
1  SELECT total
2  from invoice
3  order by total desc
4  limit 3
```

Output:

	billing_country character varying (30)	c bigint
1	USA	131
2	Canada	76
3	Brazil	61
4	France	50
5	Germany	41
6	Czech Republic	30
7	Portugal	29
8	United Kingdom	28
9	India	21
10	Chile	13
11	Ireland	13

Top 3 values of total invoice

	billing_country character varying (30)	c bigint
1	USA	131
2	Canada	76
3	Brazil	61

Questions Set 3

Which city has the best customers?

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

Query:

```
Query  Query History
1  SELECT billing_city, sum(total) as InvoiceTotal
2  from invoice
3  group by billing_city
4  order by InvoiceTotal desc
5  limit 1
```

Output:

	billing_city character varying (30) 🔒	invoicetotal double precision 🔒
1	Prague	273.24000000000007

Questions Set 3

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.

Query:

```
Query  Query History
1  v  Select c.customer_id, c.first_name, c.last_name,
2    sum(total) as s
3  from customer as c
4  join invoice as inv
5  on inv.customer_id = c.customer_id
6  group by c.customer_id
7  order by s desc
8  limit 1
```

Output:

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