

MUSIC STORE DATA ANALYSIS

Task 1- Who is the senior most employee based on job title?

Query :

```
SELECT * FROM employee
```

```
ORDER BY levels desc
```

```
LIMIT 1;
```

Output –

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	1/26/1961 0:00	1/14/2016 0:00	1008 Vrinda Ave MT	Edmonton	AB	Canada	T5K 2N1	+1 (780) 428-9482	+1 (780) 428-3457	madan.mohan@chinookcorp.com

Task 2 - Which countries have the most Invoices?

Query –

```
SELECT COUNT(*) , billing_country
```

```
FROM invoice
```

```
GROUP BY billing_country
```

```
ORDER BY count DESC
```

Output –

Count	billing_country
131	USA
76	Canada
61	Brazil
50	France
41	Germany
30	Czech Republic
29	Portugal
28	United Kingdom
21	India
13	Chile
13	Ireland
11	Spain
11	Finland
10	Australia
10	Netherlands
10	Sweden
10	Poland
10	Hungary
10	Denmark
9	Austria
9	Norway
9	Italy
7	Belgium
5	Argentina

Task 3 -What are top 3 values of total invoice?

Query –

```
SELECT total FROM invoice
```

```
ORDER BY total DESC
```

```
LIMIT 3
```

Output –

total
23.76
19.8
19.8

Task 4 – 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 –

```
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	first_name	last_name	total
5	R	Madhav	144.54

Task 5 - 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

Query –

```
SELECT sum(total)AS invoice_total, billing_city
```

FROM invoice GROUP BY billing_city

ORDER BY invoice_total DESC

Output –

invoice_total	billing_city
273.24	Prague
169.29	Mountain View
166.32	London
158.4	Berlin
151.47	Paris
129.69	SÃo Paulo
114.84	Dublin
111.87	Delhi
108.9	SÃo JosÃ© dos Campos
106.92	BrasÃlia
102.96	Lisbon
99.99	Bordeaux
99.99	MontrÃ©al
98.01	Madrid
98.01	Redmond
97.02	Santiago
94.05	Frankfurt
92.07	Orlando
91.08	Reno
91.08	Ottawa
86.13	Fort Worth
84.15	Tucson
82.17	Stuttgart
82.17	Rio de Janeiro
82.17	Porto
81.18	Sidney
79.2	New York
79.2	Edinburgh
79.2	Helsinki
78.21	Budapest
76.23	Madison
76.23	Warsaw
75.24	Yellowknife
75.24	Stockholm
73.26	Dijon
72.27	Oslo
72.27	Salt Lake City
71.28	Chicago
71.28	Bangalore
70.29	Winnipeg
69.3	Vienne
66.33	Vancouver
66.33	Boston
65.34	Amsterdam
64.35	Lyon
62.37	Halifax
60.39	Brussels
54.45	Cupertino
50.49	Rome
40.59	Toronto
39.6	Buenos Aires
37.62	Copenhagen
29.7	Edmonton

Task 6 - Write 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

Query –

```
SELECT DISTINCT email,first_name, last_name

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 email;
```

Output –

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
ftremblay@gmail.com	François	Tremblay
fzimmermann@yahoo.de	Fynn	Zimmermann
hannah.schneider@yahoo.de	Hannah	Schneider
hholy@gmail.com	Helena	Holý
hleacock@gmail.com	Heather	Leacock

hughoreilly@apple.ie	Hugh	O'Reilly
isabelle_mercier@apple.fr	Isabelle	Mercier
jacksmith@microsoft.com	Jack	Smith
jenniferp@rogers.ca	Jennifer	Peterson
jfernandes@yahoo.pt	João	Fernandes
joakim.johansson@yahoo.se	Joakim	Johansson
johavanderberg@yahoo.nl	Johannes	Van der Berg
johngordon22@yahoo.com	John	Gordon
jubarnett@gmail.com	Julia	Barnett
kachase@hotmail.com	Kathy	Chase
kara.nielsen@jubii.dk	Kara	Nielsen
ladislav_kovacs@apple.hu	Ladislav	Kovács
leonekohler@surfeu.de	Leonie	Köhler
lucas.mancini@yahoo.it	Lucas	Mancini
luisg@embraer.com.br	Luís	Gonçalves
luisrojas@yahoo.cl	Luis	Rojas
manoj.pareek@rediff.com	Manoj	Pareek
marc.dubois@hotmail.com	Marc	Dubois
mark.taylor@yahoo.au	Mark	Taylor
marthasilk@gmail.com	Martha	Silk
masampaio@sapo.pt	Madalena	Sampaio
michelleb@aol.com	Michelle	Brooks
mphilips12@shaw.ca	Mark	Philips
nschroder@surfeu.de	Niklas	Schröder
patrick.gray@aol.com	Patrick	Gray
phil.hughes@gmail.com	Phil	Hughes
puja_srivastava@yahoo.in	Puja	Srivastava
r.madhav@jetbrains.com	R	Madhav
ricunningham@hotmail.com	Richard	Cunningham
robbrown@shaw.ca	Robert	Brown
roberto.almeida@riotur.gov.br	Roberto	Almeida
stanisław.wójcik@wp.pl	Stanisław	Wójcik
steve.murray@yahoo.uk	Steve	Murray
terhi.hamalainen@apple.fi	Terhi	Hämäläinen
tgoyer@apple.com	Tim	Goyer
vstevens@yahoo.com	Victor	Stevens
wyatt.girard@yahoo.fr	Wyatt	Girard

Task 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

Query –

```
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;

Output-

Artist_id	Name	Number_of_songs
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

Task 8 –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 –

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	country	name	genre_id	rowno
17	Argentina	Alternativ	4	1
34	Australia	Rock	1	1
40	Austria	Rock	1	1
26	Belgium	Rock	1	1
205	Brazil	Rock	1	1
333	Canada	Rock	1	1
61	Chile	Rock	1	1
143	Czech Rep	Rock	1	1
24	Denmark	Rock	1	1
46	Finland	Rock	1	1
211	France	Rock	1	1
194	Germany	Rock	1	1
44	Hungary	Rock	1	1
102	India	Rock	1	1
72	Ireland	Rock	1	1
35	Italy	Rock	1	1
33	Netherlan	Rock	1	1
40	Norway	Rock	1	1
40	Poland	Rock	1	1
108	Portugal	Rock	1	1
46	Spain	Rock	1	1
60	Sweden	Rock	1	1
166	United Kir	Rock	1	1
561	USA	Rock	1	1

Task 9 – 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 –

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 Customer_with_country WHERE RowNo <= 1

```

Output-

customer_id	first_name	last_name	billing_country	total_spending	rowno
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	60.39	1
1	Luís	Gonçalves	Brazil	108.9	1
3	François	Tremblay	Canada	99.99	1
57	Luis	Rojas	Chile	97.02	1
5	R	Madhav	Czech Republic	144.54	1
9	Kara	Nielsen	Denmark	37.62	1
44	Terhi	Hämäläinen	Finland	79.2	1
42	Wyatt	Girard	France	99.99	1
37	Fynn	Zimmermann	Germany	94.05	1
45	Ladislav	Kovács	Hungary	78.21	1
58	Manoj	Pareek	India	111.87	1
46	Hugh	O'Reilly	Ireland	114.84	1
47	Lucas	Mancini	Italy	50.49	1
48	Johannes	Van der Berg	Netherlands	65.34	1
4	Björn	Hansen	Norway	72.27	1
49	Stanisław	Wójcik	Poland	76.23	1
34	João	Fernandes	Portugal	102.96	1
50	Enrique	Muñoz	Spain	98.01	1
51	Joakim	Johansson	Sweden	75.24	1
53	Phil	Hughes	United Kingdom	98.01	1
17	Jack	Smith	USA	98.01	1

Task 10 – 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 alb ON alb.album_id = t.album_id

JOIN best_selling_artist bsa ON bsa.artist_id = alb.artist_id

GROUP BY 1,2,3,4

ORDER BY 5 DESC;

```

Output-

Customer_id	First_name	Last_name	Artist_name	Amount_spent
46	Hugh	O'Reilly	Queen	27.72
38	Niklas	Schröder	Queen	18.81
3	François	Tremblay	Queen	17.82
34	João	Fernandes	Queen	16.83
53	Phil	Hughes	Queen	11.88
41	Marc	Dubois	Queen	11.88
47	Lucas	Mancini	Queen	10.89
33	Ellie	Sullivan	Queen	10.89
20	Dan	Miller	Queen	3.96
5	R	Madhav	Queen	3.96
23	John	Gordon	Queen	2.97
54	Steve	Murray	Queen	2.97
31	Martha	Silk	Queen	2.97
16	Frank	Harris	Queen	1.98
17	Jack	Smith	Queen	1.98
24	Frank	Ralston	Queen	1.98
30	Edward	Francis	Queen	1.98
35	Madalena	Sampaio	Queen	1.98
36	Hannah	Schneider	Queen	1.98
11	Alexandre	Rocha	Queen	1.98
8	Daan	Peeters	Queen	1.98
42	Wyatt	Girard	Queen	1.98
44	Terhi	Hämäläinen	Queen	1.98
1	Luís	Gonçalves	Queen	1.98
48	Johannes	Van der Berg	Queen	1.98
49	Stanisław	Wójcik	Queen	1.98
52	Emma	Jones	Queen	1.98
57	Luis	Rojas	Queen	1.98
15	Jennifer	Peterson	Queen	1.98
28	Julia	Barnett	Queen	1.98
27	Patrick	Gray	Queen	0.99
58	Manoj	Pareek	Queen	0.99
45	Ladislav	Kovács	Queen	0.99
26	Richard	Cunningham	Queen	0.99
59	Puja	Srivastava	Queen	0.99
13	Fernanda	Ramos	Queen	0.99
6	Helena	Holmér	Queen	0.99
22	Heather	Leacock	Queen	0.99
19	Tim	Goyer	Queen	0.99
39	Camille	Bernard	Queen	0.99
55	Mark	Taylor	Queen	0.99
50	Enrique	Muñoz	Queen	0.99
43	Isabelle	Mercier	Queen	0.99