

Project: Music Store

Michał Lubczyński

a) Provide a list of artists along with the number of genres they play. List sorted in descending order by musical genre and ascending by artist name. Additionally, in the third column give the place in dense rank in descending order. Limit results to 20 records.

```
SELECT DISTINCT a.name "Performer_name",
               count(DISTINCT g.genreid) "Music_genre_count",
               DENSE_RANK() over(ORDER BY count(DISTINCT g.genreid) DESC) "Rank"
FROM artist a
JOIN album al ON a.artistid = al.artistid
JOIN track t ON al.albumid = t.albumid
JOIN genre g ON t.genreid = g.genreid
GROUP BY a.name
ORDER BY "Music_genre_count" DESC
LIMIT 20
```

| Performer_name | Music_genre_count | Rank |
|-----------------------|-------------------|------|
| Iron Maiden | 4 | 1 |
| Audioslave | 3 | 2 |
| Battlestar Galactica | 3 | 2 |
| Gilberto Gil | 3 | 2 |
| Jamiroquai | 3 | 2 |
| Lenny Kravitz | 3 | 2 |
| Various Artists | 3 | 2 |
| Amy Winehouse | 2 | 3 |
| Antônio Carlos Jobim | 2 | 3 |
| Eric Clapton | 2 | 3 |
| Faith No More | 2 | 3 |
| Foo Fighters | 2 | 3 |
| Guns N' Roses | 2 | 3 |
| Heroes | 2 | 3 |
| Lost | 2 | 3 |
| Ozzy Osbourne | 2 | 3 |
| Pearl Jam | 2 | 3 |
| Red Hot Chili Peppers | 2 | 3 |
| R.E.M. | 2 | 3 |
| The Office | 2 | 3 |

b) What kinds of music genres do Canadian customers listen to? Display genre and number of distinct songs in the genre.

```
SELECT DISTINCT g.name "Genre",
               count(DISTINCT t.trackid) "Song_count"
```

```

FROM genre g
JOIN track t ON g.genreid = t.genreid
JOIN invoiceline i ON t.trackid = i.trackid
JOIN invoice inv ON i.invoiceid = inv.invoiceid
WHERE inv.billingcountry = 'Canada'
GROUP BY "Genre"
ORDER BY "Song_count" DESC

```

| Genre | Song_count |
|--------------------|------------|
| Rock | 106 |
| Latin | 60 |
| Metal | 39 |
| Alternative & Punk | 36 |
| Jazz | 13 |
| Bossa Nova | 7 |
| Reggae | 7 |
| World | 6 |
| Classical | 5 |
| Hip Hop/Rap | 5 |
| R&B/Soul | 5 |
| Blues | 4 |
| Electronica/Dance | 4 |
| Drama | 2 |
| Rock And Roll | 2 |
| TV Shows | 1 |

c) List the ten customers who spent the most in this store.

```

SELECT  firstname,
        lastname,
        sum(total)
FROM invoice
JOIN customer ON invoice.customerid = customer.customerid
GROUP BY invoice.customerid, firstname, lastname
ORDER BY sum(total) DESC
LIMIT 10

```

| firstname | lastname | sum |
|-----------|------------|-------|
| Helena | Holý | 49.62 |
| Richard | Cunningham | 47.62 |
| Luis | Rojas | 46.62 |
| Ladislav | Kovács | 45.62 |
| Hugh | O'Reilly | 45.62 |
| Frank | Ralston | 43.62 |
| Julia | Barnett | 43.62 |
| Fynn | Zimmermann | 43.62 |
| Astrid | Gruber | 42.62 |
| Victor | Stevens | 42.62 |

d) Give the distribution of the sum of money spent with the breakdown by customer countries as a percentage with an accuracy of one per mille. Sort the result by the largest share.

```
SELECT billingcountry,
       round(sum(total)*100/(SELECT sum(total) FROM invoice), 1) "%_share"
FROM invoice
GROUP BY billingcountry
ORDER BY sum(total) DESC
```

| billingcountry | %_share |
|----------------|---------|
| USA | 22.5 |
| Canada | 13.1 |
| France | 8.4 |
| Brazil | 8.2 |
| Germany | 6.7 |
| United Kingdom | 4.8 |
| Czech Republic | 3.9 |
| Portugal | 3.3 |
| India | 3.2 |
| Chile | 2.0 |
| Ireland | 2.0 |
| Hungary | 2.0 |
| Austria | 1.8 |
| Finland | 1.8 |
| Netherlands | 1.7 |
| Norway | 1.7 |
| Sweden | 1.7 |
| Argentina | 1.6 |
| Belgium | 1.6 |
| Poland | 1.6 |
| Australia | 1.6 |
| Italy | 1.6 |
| Denmark | 1.6 |
| Spain | 1.6 |

e) Display the percentage (with an accuracy of two decimal places) of the types of music file formats purchased from the entire data set (and separately by genre). What genre of music has nobody bought?

```
WITH cte_counter AS
  (SELECT gen.name genre_name,
         med.name format_name,
         count(inv.trackid) OVER(PARTITION BY med.name) format_song_count,
         round((count(inv.trackid) OVER
                (PARTITION BY med.name)*100.0 /
                (SELECT count(trackid) FROM invoiceline))), 2) "%_format",
         count(inv.trackid) over(PARTITION BY gen.name, med.name) genre_song_count,
         round((count(inv.trackid) OVER
                (PARTITION BY gen.name, med.name)*100.0 /
                (SELECT count(trackid) FROM invoiceline))), 2) "%_genre"
  FROM invoiceline AS inv
  JOIN track t ON inv.trackid = t.trackid
```

```

JOIN mediatype med ON t.mediatypeid = med.mediatypeid
RIGHT JOIN genre gen ON t.genreid = gen.genreid)
SELECT DISTINCT *
FROM cte_counter
ORDER BY format_song_count DESC , genre_song_count DESC

```

| genre_name | format_name | format_song_count | %_format | genre_song_count | %_genre |
|-------------------|---------------|-------------------|----------|------------------|---------|
| Rock | MPEG audio | 1976 | 88.21 | 773 | 34.51 |
| Latin | file | | | | |
| | MPEG audio | 1976 | 88.21 | 385 | 17.19 |
| Metal | file | | | | |
| | MPEG audio | 1976 | 88.21 | 264 | 11.79 |
| Alternative & | file | | | | |
| Punk | MPEG audio | 1976 | 88.21 | 244 | 10.89 |
| Jazz | file | | | | |
| | MPEG audio | 1976 | 88.21 | 80 | 3.57 |
| Blues | file | | | | |
| | MPEG audio | 1976 | 88.21 | 61 | 2.72 |
| Reggae | file | | | | |
| | MPEG audio | 1976 | 88.21 | 30 | 1.34 |
| R&B/Soul | file | | | | |
| | MPEG audio | 1976 | 88.21 | 29 | 1.29 |
| Soundtrack | file | | | | |
| | MPEG audio | 1976 | 88.21 | 20 | 0.89 |
| Hip Hop/Rap | file | | | | |
| | MPEG audio | 1976 | 88.21 | 17 | 0.76 |
| Bossa Nova | file | | | | |
| | MPEG audio | 1976 | 88.21 | 15 | 0.67 |
| Heavy Metal | file | | | | |
| | MPEG audio | 1976 | 88.21 | 12 | 0.54 |
| World | file | | | | |
| | MPEG audio | 1976 | 88.21 | 12 | 0.54 |
| Electronica/Dance | file | | | | |
| | MPEG audio | 1976 | 88.21 | 12 | 0.54 |
| Easy Listening | file | | | | |
| | MPEG audio | 1976 | 88.21 | 10 | 0.45 |
| Rock And Roll | file | | | | |
| | MPEG audio | 1976 | 88.21 | 6 | 0.27 |
| Pop | file | | | | |
| | MPEG audio | 1976 | 88.21 | 6 | 0.27 |
| Rock | file | | | | |
| | Protected AAC | 146 | 6.52 | 61 | 2.72 |
| Classical | audio file | | | | |
| | Protected AAC | 146 | 6.52 | 37 | 1.65 |
| Pop | audio file | | | | |
| | Protected AAC | 146 | 6.52 | 22 | 0.98 |
| Alternative | audio file | | | | |
| | Protected AAC | 146 | 6.52 | 14 | 0.63 |
| R&B/Soul | audio file | | | | |
| | Protected AAC | 146 | 6.52 | 12 | 0.54 |
| TV Shows | audio file | | | | |
| | Protected | 111 | 4.96 | 47 | 2.10 |
| | MPEG-4 video | | | | |
| Drama | file | | | | |
| | Protected | 111 | 4.96 | 29 | 1.29 |
| | MPEG-4 video | | | | |
| Sci Fi & Fantasy | file | | | | |
| | Protected | 111 | 4.96 | 20 | 0.89 |
| | MPEG-4 video | | | | |
| Comedy | file | | | | |
| | Protected | 111 | 4.96 | 9 | 0.40 |
| | MPEG-4 video | | | | |
| | file | | | | |

| | | | | | |
|-----------------|-----------------------------------|-----|------|---|------|
| Science Fiction | Protected MPEG-4 video file | 111 | 4.96 | 6 | 0.27 |
| Classical | Purchased AAC audio file | 4 | 0.18 | 4 | 0.18 |
| Latin | AAC audio file | 3 | 0.13 | 1 | 0.04 |
| World | AAC audio file | 3 | 0.13 | 1 | 0.04 |
| Rock | AAC audio file | 3 | 0.13 | 1 | 0.04 |
| Opera | NA | 0 | 0.00 | 0 | 0.00 |

f) What artist is most often bought by those who also bought Miles Davis' albums (excluding Miles Davis and Various Artists)

```
SELECT art.name,
       count(*) "song_count"
FROM artist art
  JOIN album al ON art.artistid = al.artistid
  JOIN track tr ON al.albumid = tr.albumid
  JOIN invoiceline inv ON tr.trackid = inv.trackid
  JOIN invoice i ON inv.invoiceid = i.invoiceid
  JOIN customer cus ON i.customerid = cus.customerid
WHERE cus.customerid IN (SELECT cus.customerid "Miles_fan"
                        FROM artist art
                          JOIN album al ON art.artistid = al.artistid
                          JOIN track tr ON al.albumid = tr.albumid
                          JOIN invoiceline inv ON tr.trackid = inv.trackid
                          JOIN invoice i ON inv.invoiceid = i.invoiceid
                          JOIN customer cus ON i.customerid = cus.customerid
                        WHERE art.name LIKE 'Miles%')
      AND art.name NOT LIKE 'Miles%'
      AND art.name NOT LIKE 'Various%'
GROUP BY art.name
ORDER BY count(*) DESC
LIMIT 1
```

| name | song_count |
|-----------------------|------------|
| Red Hot Chili Peppers | 15 |

g) Create a ranking of customer service representatives (supportrepid), where the evaluation criterion is the obtained highest non-zero turnover for a given month. The query should return a table with the name and surname of the employee of the month for each sales month.

```
WITH cte AS
    (SELECT EXTRACT(YEAR FROM invoicedate) "Year",
        EXTRACT(MONTH FROM invoicedate) "Month",
        supportrepid "Employee_id",
        employee.firstname "Employee_firstname",
        employee.lastname "Employee_lastname",
        sum(total) "Sale_sum",
        rank() OVER(PARTITION BY EXTRACT(YEAR FROM invoicedate),
            EXTRACT(MONTH FROM invoicedate)
            ORDER BY sum(total) DESC) "Emp_Rank"
    FROM invoice
        JOIN customer ON invoice.customerid = customer.customerid
        JOIN employee ON customer.supportrepid = employee.employeeid
    GROUP BY "Year", "Month", "Employee_id",
        "Employee_firstname", "Employee_lastname"
    ORDER BY "Year", "Month", "Sale_sum" DESC)
SELECT "Year",
    "Month",
    "Employee_id",
    "Employee_firstname",
    "Employee_lastname",
    max("Sale_sum")
FROM cte
WHERE "Emp_Rank" = 1
GROUP BY "Year", "Month", "Employee_id",
    "Employee_firstname", "Employee_lastname"
ORDER BY "Year", "Month"
```

| Year | Month | Employee_id | Employee_firstname | Employee_lastname | max |
|------|-------|-------------|--------------------|-------------------|-------|
| 2007 | 1 | 4 | Margaret | Park | 23.76 |
| 2007 | 2 | 3 | Jane | Peacock | 20.79 |
| 2007 | 3 | 5 | Steve | Johnson | 21.78 |
| 2007 | 4 | 3 | Jane | Peacock | 18.81 |
| 2007 | 5 | 5 | Steve | Johnson | 24.75 |
| 2007 | 6 | 5 | Steve | Johnson | 24.75 |
| 2007 | 7 | 3 | Jane | Peacock | 22.77 |
| 2007 | 8 | 3 | Jane | Peacock | 30.69 |
| 2007 | 9 | 4 | Margaret | Park | 28.71 |
| 2007 | 10 | 5 | Steve | Johnson | 29.70 |
| 2007 | 11 | 4 | Margaret | Park | 31.68 |
| 2007 | 12 | 5 | Steve | Johnson | 15.84 |
| 2008 | 1 | 5 | Steve | Johnson | 48.66 |
| 2008 | 2 | 3 | Jane | Peacock | 31.77 |
| 2008 | 3 | 3 | Jane | Peacock | 34.72 |
| 2008 | 4 | 3 | Jane | Peacock | 26.73 |
| 2008 | 5 | 4 | Margaret | Park | 21.78 |
| 2008 | 6 | 4 | Margaret | Park | 22.77 |
| 2008 | 7 | 3 | Jane | Peacock | 21.78 |
| 2008 | 8 | 3 | Jane | Peacock | 17.82 |

| | | | | | |
|------|----|---|----------|---------|-------|
| 2008 | 9 | 4 | Margaret | Park | 17.82 |
| 2008 | 10 | 5 | Steve | Johnson | 15.84 |
| 2008 | 11 | 3 | Jane | Peacock | 30.69 |
| 2008 | 12 | 3 | Jane | Peacock | 26.73 |
| 2009 | 1 | 5 | Steve | Johnson | 26.73 |
| 2009 | 2 | 4 | Margaret | Park | 15.84 |
| 2009 | 3 | 3 | Jane | Peacock | 19.80 |
| 2009 | 4 | 3 | Jane | Peacock | 36.77 |
| 2009 | 5 | 5 | Steve | Johnson | 24.80 |
| 2009 | 6 | 4 | Margaret | Park | 20.83 |
| 2009 | 7 | 3 | Jane | Peacock | 25.74 |
| 2009 | 8 | 5 | Steve | Johnson | 25.74 |
| 2009 | 9 | 3 | Jane | Peacock | 21.78 |
| 2009 | 10 | 3 | Jane | Peacock | 28.71 |
| 2009 | 11 | 5 | Steve | Johnson | 23.76 |
| 2009 | 12 | 4 | Margaret | Park | 23.76 |
| 2010 | 1 | 4 | Margaret | Park | 16.83 |
| 2010 | 2 | 4 | Margaret | Park | 10.89 |
| 2010 | 3 | 4 | Margaret | Park | 18.81 |
| 2010 | 4 | 5 | Steve | Johnson | 29.70 |
| 2010 | 5 | 3 | Jane | Peacock | 31.68 |
| 2010 | 6 | 4 | Margaret | Park | 22.77 |
| 2010 | 7 | 5 | Steve | Johnson | 35.66 |
| 2010 | 8 | 4 | Margaret | Park | 39.70 |
| 2010 | 9 | 4 | Margaret | Park | 24.82 |
| 2010 | 10 | 3 | Jane | Peacock | 24.78 |
| 2010 | 11 | 4 | Margaret | Park | 24.75 |
| 2010 | 12 | 3 | Jane | Peacock | 22.77 |
| 2011 | 1 | 3 | Jane | Peacock | 19.80 |
| 2011 | 2 | 3 | Jane | Peacock | 15.84 |
| 2011 | 3 | 5 | Steve | Johnson | 17.82 |
| 2011 | 4 | 4 | Margaret | Park | 33.66 |
| 2011 | 5 | 5 | Steve | Johnson | 20.79 |
| 2011 | 6 | 3 | Jane | Peacock | 34.65 |
| 2011 | 7 | 4 | Margaret | Park | 18.81 |
| 2011 | 8 | 4 | Margaret | Park | 13.86 |
| 2011 | 9 | 5 | Steve | Johnson | 15.84 |
| 2011 | 10 | 4 | Margaret | Park | 19.80 |
| 2011 | 11 | 5 | Steve | Johnson | 31.80 |
| 2011 | 12 | 3 | Jane | Peacock | 21.79 |

h) For each of the three employees (supportrepid), indicate the months for which they recorded zero sales.

```
WITH cte AS
    (SELECT cus.supportrepid "employee",
        EXTRACT (YEAR FROM inv.invoicedate) "year",
        EXTRACT (MONTH FROM inv.invoicedate) "month"
    FROM customer cus
        JOIN invoice inv ON cus.customerid = inv.customerid
    ORDER BY "employee", "year", "month")
```

```

SELECT sub1.employee,
       sub1.year,
       CASE WHEN sub1.month = 12 THEN 1
            ELSE sub1.month+1 END AS "month"
FROM (SELECT "employee",
            "year",
            "month",
            lead("month") OVER (ORDER BY "employee") "lead1"
      FROM cte) "sub1"
WHERE sub1.lead1 - sub1.month > 1 OR sub1.lead1 - sub1.month = -10
GROUP BY sub1.employee, sub1.year, sub1.month, sub1.lead1
ORDER BY sub1.employee, sub1.year, sub1.month

```

| employee | year | month |
|----------|------|-------|
| 3 | 2007 | 10 |
| 3 | 2011 | 4 |
| 4 | 2007 | 1 |
| 4 | 2011 | 6 |
| 5 | 2007 | 8 |
| 5 | 2008 | 3 |
| 5 | 2008 | 7 |
| 5 | 2011 | 2 |
| 5 | 2011 | 4 |