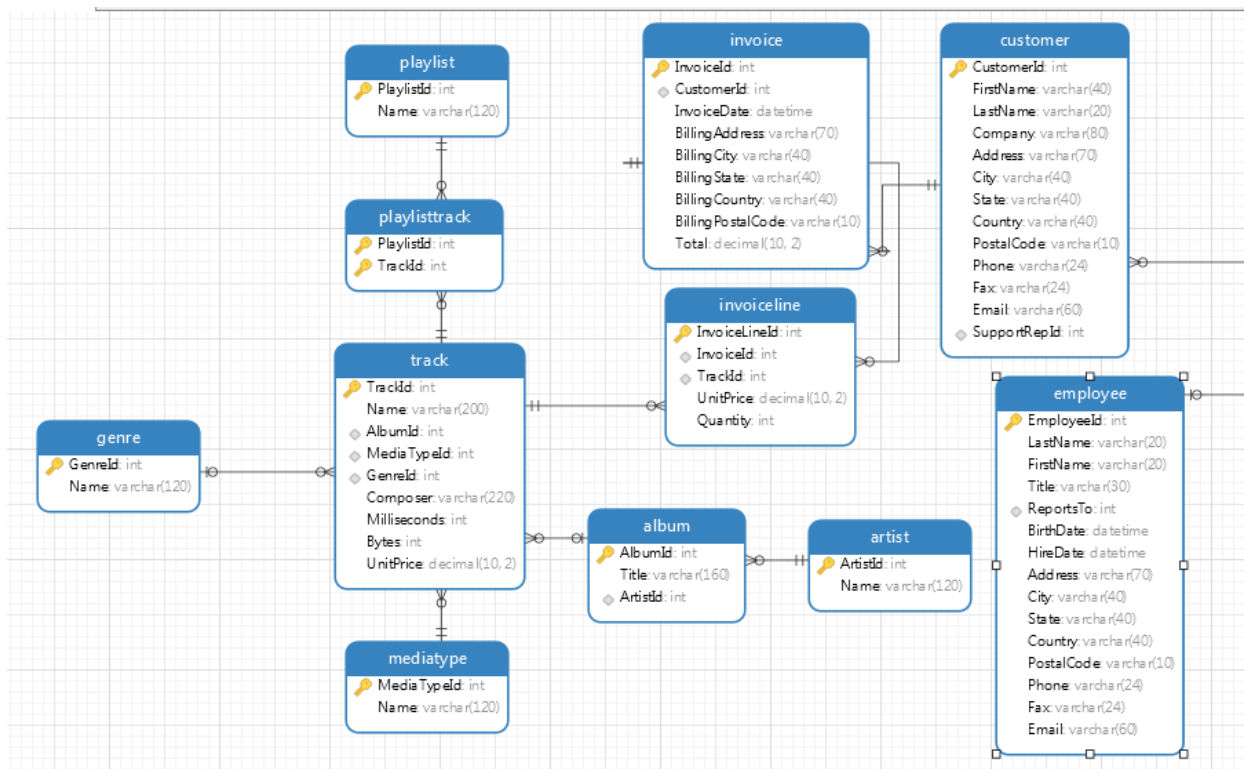


رسم نمودار ER و تشخیص جداول fact و dimension

- نمودار ER جداول پایگاه داده به صورت زیر است:



در طرح داده پایگاه داده Chinook، می‌توانیم جداول fact و dimension را بر اساس مدل داده‌ی فوق که نمایانگر یک فروشگاه رسانه دیجیتال است، شناسایی کنیم. با توجه به ساختار نمایش داده شده در شکل بالا، این اسکما به اسکمای star نزدیک تر است. یک جدول fact مرکزی داریم که در این اسکما، جدول fact InvoiceLine می‌باشد، که توسط جداول dimension احاطه شده است. جدول fact جزئیات هر فروش را ثبت می‌کند، شامل:

- `InvoiceLineId` (کلید اصلی)

- `InvoiceId` (کلید خارجی به فاکتور)

- `TrackId` (کلید خارجی به آهنگ)

- `UnitPrice`: قیمت آهنگ در زمان فروش.

- `Quantity`: تعداد واحدهای فروخته شده.

جداول dimension نیز شامل ویژگی‌های توصیفی مرتبط با factها هستند و معمولاً نرمالیزه شده‌اند.
جداول dimension کلیدی در پایگاه داده Chinook شامل موارد زیر است:

- Customers: اطلاعات مشتریان را شامل می‌شود.

- `CustomerId` (کلید اصلی)

- `FirstName`

- `LastName`

- `Country`

- `Email`

- Tracks: اطلاعات مربوط به آهنگ‌ها را شامل می‌شود.

- `TrackId` (کلید اصلی)

- `Name`

- `AlbumId` (کلید خارجی به آلبوم)

- `GenreId` (کلید خارجی به ژانر)

- `Composer`

- Albums: جزئیات آلبوم‌ها را شامل می‌شود.

- `AlbumId` (کلید اصلی)

- `Title`

- `ArtistId` (کلید خارجی به هنرمند)

– Artists: اطلاعات مربوط به هنرمندان را شامل می‌شود.

– `ArtistId` (کلید اصلی)

– `Name`

– Genres: اطلاعات مربوط به ژانرها را شامل می‌شود.

– `GenreId` (کلید اصلی)

– `Name`

– Invoices: جزئیات فاکتورها را شامل می‌شود.

– `InvoiceId` (کلید اصلی)

– `CustomerId` (کلید خارجی به مشتری)

– `InvoiceDate`

– `BillingAddress`

– `BillingCity`

– `BillingState`

– `BillingCountry`

– `BillingPostalCode`

بنابراین، جدول InvoiceLine به عنوان جدول فکت عمل می‌کند و داده‌های تراکنشی را ثبت می‌کند، در حالی که جداول Customers, Tracks, Albums, Artists, Genres و Invoices به عنوان جداول بعد عمل می‌کنند و زمینه‌ای برای داده‌های فروش فراهم می‌آورند. این ساختار از تحلیل فروش بر اساس ابعاد مختلف، مانند ویژگی‌های جمعیت‌شناختی مشتریان، جزئیات آهنگ‌ها و طبقه‌بندی‌های ژانر پشتیبانی می‌کند.

بخش کوئری های SQL

1- 10 آهنگ برتر که بیشترین درآمد رو داشتن به همراه درآمد ایجاد شده

```
SELECT |
    t.Name AS TrackName,
    SUM(il.Quantity * il.UnitPrice) AS Income
FROM
    InvoiceLine il
JOIN
    Track t ON il.TrackId = t.TrackId
GROUP BY
    t.TrackId
ORDER BY
    Income DESC
LIMIT 10;
```

خروجی کوئری:

TrackName	Income
Hot Girl	3.98
Walkabout	3.98
Gay Witch Hunt	3.98
How to Stop an Exploding	3.98
Pilot	3.98
Phyllis's Wedding	3.98
The Fix	3.98
The Woman King	3.98
I Do	1.99
The Glass Ballerina	1.99

2- محبوب ترین ژانر، به ترتیب از نظر تعداد آهنگهای فروخته شده و کل درآمد

```
SELECT
    subquery.GenreName, COUNT(il.TrackId) AS TotalSoldTrack, SUM(il.Quantity * il.UnitPrice) AS Income
FROM
    (SELECT
        t.TrackId, g.`Name` AS GenreName
    FROM
        Track t
    JOIN
        Genre g ON t.GenreId = g.GenreId
    GROUP BY
        t.TrackId, g.GenreId, GenreName) AS subquery
JOIN
    invoiceLine il ON subquery.TrackId = il.TrackId
GROUP BY
    subquery.GenreName
ORDER BY
    TotalSoldTrack DESC, income DESC
LIMIT 1;
```

خروجی کوئری:

GenreName	TotalSoldTrack	Income
Rock	835	826.65

3- کاربرانی که تا حالا خرید نداشتند.

```
SELECT
    c.CustomerId,
    c.FirstName,
    c.LastName,
    c.Email
FROM
    Customer c
LEFT JOIN
    Invoice i ON c.CustomerId = i.CustomerId
WHERE
    i.InvoiceId IS NULL;
```

خروجی کوئری:

CustomerId	FirstName	LastName	Email
(N/A)	(N/A)	(N/A)	(N/A)

4- میانگین زمان آهنگ ها در در هر آلبوم

```
SELECT
    al.Title, AVG(t.Milliseconds)/1000 AS AveragOfTrackDurationSeconds
FROM
    Album al
JOIN
    Track t ON t.AlbumId = al.AlbumId
GROUP BY
    al.Title, al.AlbumId
```

خروجی کوئری:

Title	AveragOfTrackDurationSeconds
For Those About To Rock \	240.04150000
Balls to the Wall	342.56200000
Restless and Wild	286.02933333
Let There Be Rock	306.65737500
Big Ones	294.11393333
Jagged Little Pill	265.45576923
Facelift	270.78041667
Warner 25 Anos	207.63757143
Plays Metallica By Four Cel	333.92587500
Audioslave	280.55092857
Out Of Exile	268.68641667
BackBeat Soundtrack	134.64350000

5- کارمندی که بیشترین تعداد فروش را داشته

```

SELECT
    e.FirstName,
    e.LastName,
    COUNT(i.InvoiceId) AS TotalSales
FROM
    Employee e
JOIN
    Customer c ON e.EmployeeId = c.SupportRepId
JOIN
    Invoice i ON c.CustomerId = i.CustomerId
GROUP BY
    e.EmployeeId
ORDER BY
    TotalSales DESC
LIMIT 1;

```

خروجی کوئری:

FirstName	LastName	TotalSales
▶ Jane	Peacock	146

6- کاربرانی که از بیش از یک ژانر خرید کردند.

```

WITH CustomerGenres AS (
    SELECT
        c.CustomerId,
        c.FirstName,
        c.LastName,
        g.Name AS GenreName
    FROM
        Customer c
    JOIN
        Invoice i ON c.CustomerId = i.CustomerId
    JOIN
        Invoiceline il ON i.InvoiceId = il.InvoiceId
    JOIN
        Track t ON il.TrackId = t.TrackId
    JOIN
        Genre g ON t.GenreId = g.GenreId
)
SELECT
    CustomerId,
    FirstName,
    LastName,
    COUNT(DISTINCT GenreName) AS NumGenres
FROM
    CustomerGenres
GROUP BY
    CustomerId,
    FirstName,
    LastName
HAVING
    COUNT(DISTINCT GenreName) > 1

```

خروجی کوئری:

CustomerId	FirstName	LastName	NumGenres
1	Luís	Gonçalves	8
2	Leonie	Köhler	7
3	François	Tremblay	10
4	Bjørn	Hansen	8
5	František	Wichterlová	8
6	Helena	Holý	9
7	Astrid	Gruber	9
8	Daan	Peeters	4
9	Kara	Nielsen	5
10	Eduardo	Martins	7
11	Alexandre	Rocha	6
12	Roberto	Almeida	5
13	Fernanda	Ramos	7
14	Mark	Philips	10
15	Jennifer	Peterson	8

7- سه آهنگ برتر از نظر درآمد فروش برای هر ژانر

```

WITH GenreSales AS (
    SELECT
        g.Name AS Genre,
        t.Name AS Track,
        SUM(il.Quantity * il.UnitPrice) AS TotalSales,
        ROW_NUMBER() OVER (PARTITION BY g.Name ORDER BY SUM(il.Quantity * il.UnitPrice) DESC) AS RowNum
    FROM InvoiceLine il
    JOIN Track t ON il.TrackId = t.TrackId
    JOIN Genre g ON t.GenreId = g.GenreId
    GROUP BY g.Name, t.Name
)
SELECT
    Genre,
    Track,
    TotalSales
FROM GenreSales
WHERE RowNum <= 3
ORDER BY Genre, TotalSales DESC;

```

خروجی کوئری:

Genre	Track	TotalSales
Alternative	Say Hello 2 Heaven	0.99
Alternative	Call Me a Dog	0.99
Alternative	Four Walled World	0.99
Alternative & Punk	Release	2.97
Alternative & Punk	Radio Song	1.98
Alternative & Punk	Take It Or Leave It	1.98
Blues	Travis Walk	1.98
Blues	Midnight From The Inside	1.98
Blues	Title Song	1.98
Bossa Nova	Onde Anda Você	1.98
Bossa Nova	Pot-Pourri N.º 4	0.99
Bossa Nova	Samba Da Volta	0.99
Classical	Scheherazade, Op. 35: I. Th	1.98
Classical	Symphonie Fantastique, O	1.98
Classical	Suite No. 3 in D, BWV 1068	1.98

8- تعداد آهنگهای فروخته شده به صورت تجمعی در هر سال به صورت جداگانه.

```
SELECT
    YEAR(i.InvoiceDate) AS Year,
    SUM(il.Quantity) AS TotalTracksSold
FROM
    InvoiceLine il
JOIN
    Invoice i ON il.InvoiceId = i.InvoiceId
GROUP BY
    Year
ORDER BY
    Year;
```

خروجی کوئری:

Year	TotalTracksSold
2021	454
2022	455
2023	442
2024	447

9- کاربرانی که مجموع خریدشان بالاتر از میانگین مجموع خرید تمام کاربران است.


```

WITH CustomerSale AS(
    SELECT
        c.LastName,
        SUM(il.UnitPrice*il.Quantity) as income
    FROM
        Invoice i
    JOIN
        InvoiceLine il ON i.InvoiceId=il.InvoiceId
    JOIN
        Customer c ON c.CustomerId=i.CustomerId
    GROUP BY
        c.LastName)

SELECT
    LastName, income
FROM
    CustomerSale
HAVING
    income > (SELECT AVG(income) from CustomerSale)
ORDER BY
    income DESC;

```

خروجی کوئری:

LastName	income
Holý	49.62
Cunningham	47.62
Rojas	46.62
Kovács	45.62