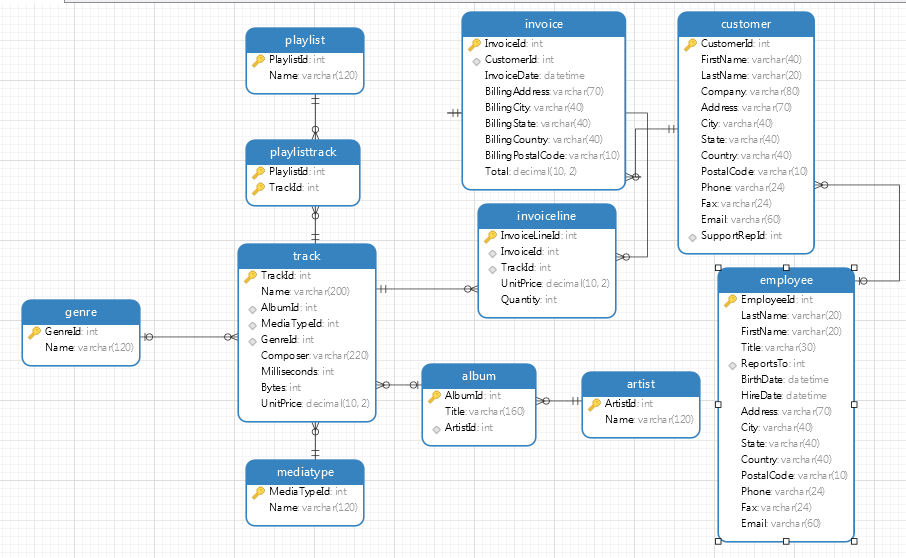
**فاز اول پروژه تحلیل داده زهرا محمودآبادی، تابستان 1403**

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 جدول InvoiceLine می باشد، 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`

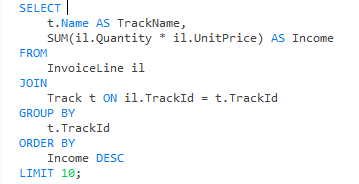
- `BillingCountry`

- `BillingPostalCode`

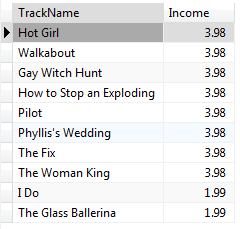
بنابراین، جدول InvoiceLine به عنوان جدول فکت عمل می‌کند و داده‌های تراکنشی را ثبت می‌کند، در حالی که جداول Customers، Tracks، Albums، Artists، Genres و Invoices به عنوان جداول بعد عمل می‌کنند و زمینه‌ای برای داده‌های فروش فراهم می‌آورند. این ساختار از تحلیل فروش بر اساس ابعاد مختلف، مانند ویژگی‌های جمعیت‌شناختی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

بخش کوئری های **SQL**

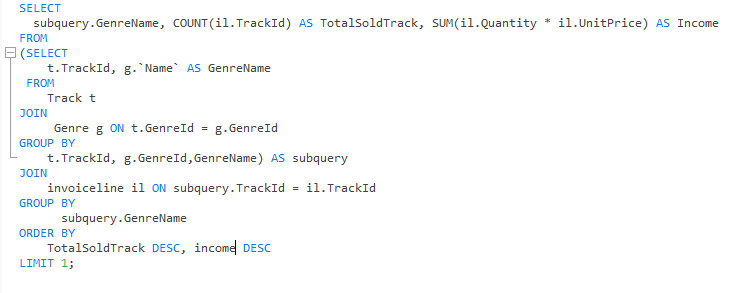
1. 10 آهنگ برتر که بیشترین درامد رو داشتن به همراه درامد ایجاد شده



خروجی کوئری:



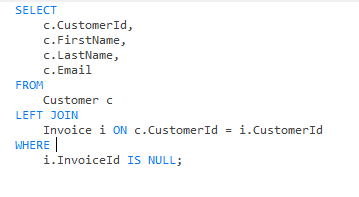
1. محبوب ترین ژانر، به ترتیب از نظر تعداد آهنگهای فروخته شده و کل درآمد



خروجی کوئری:



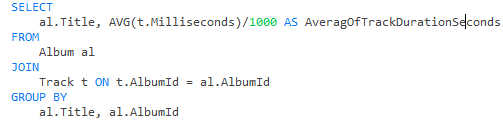
1. کاربرانی که تا حالا خرید نداشتند.



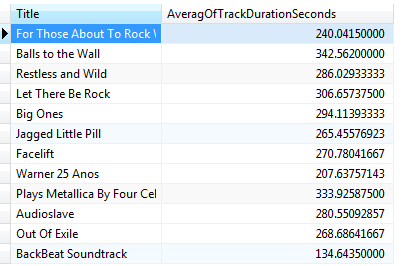
خروجی کوئری:



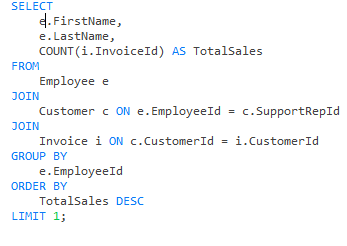
1. میانگین زمان آهنگ ها در در هر آلبوم



خروجی کوئری:



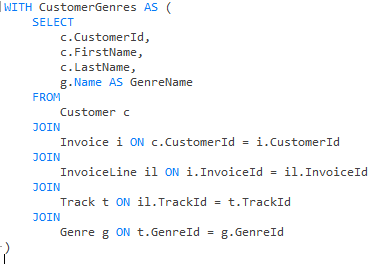
1. کارمندی که بیشترین تعداد فروش را داشته

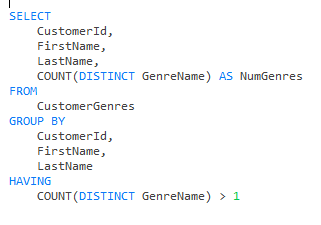


خروجی کوئری:

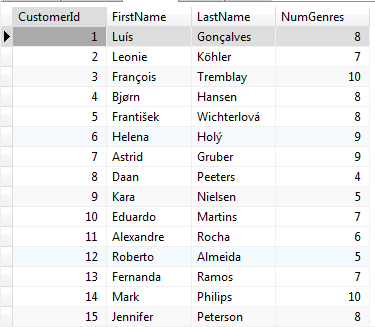


1. کاربرانی که از بیش از یک ژانر خرید کردند.

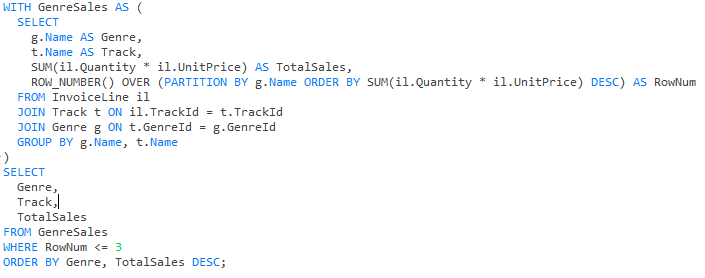




خروجی کوئری:



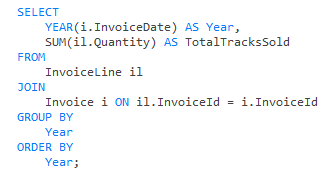
1. سه آهنگ برتر از نظر درامد فروش برای هر ژانر



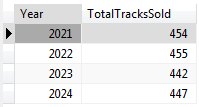
خروجی کوئری:



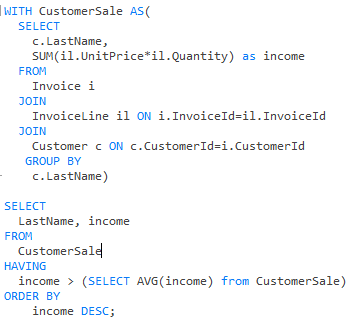
1. تعداد آهنگهای فروخته شده به صورت تجمعی در هر سال به صورت جداگانه.



خروجی کوئری:



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



خروجی کوئری:

