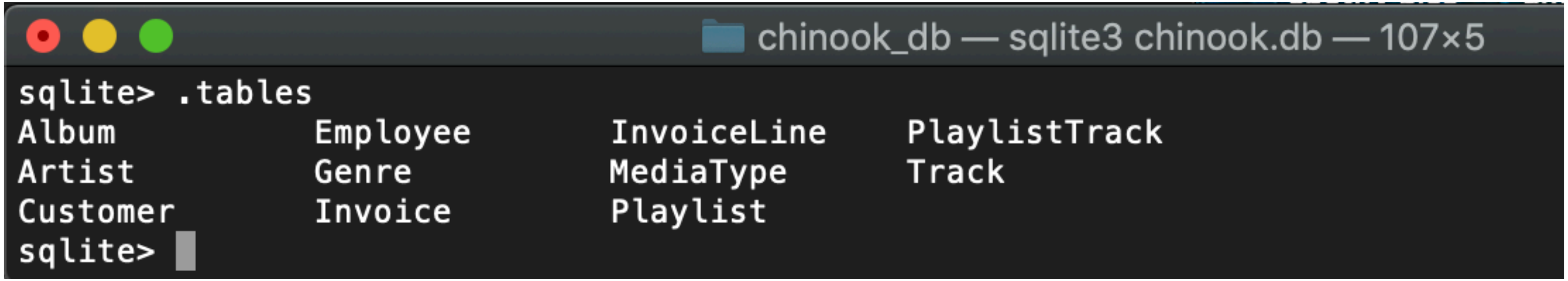
SQL FOR DATA ANALYSIS PROJECT FROM UDACITY

- The database used (chinook.db) represents a digital media store, including tables for artists, albums, media tracks, invoices and customers

- I used sqlite3 via the Terminal App to query the database

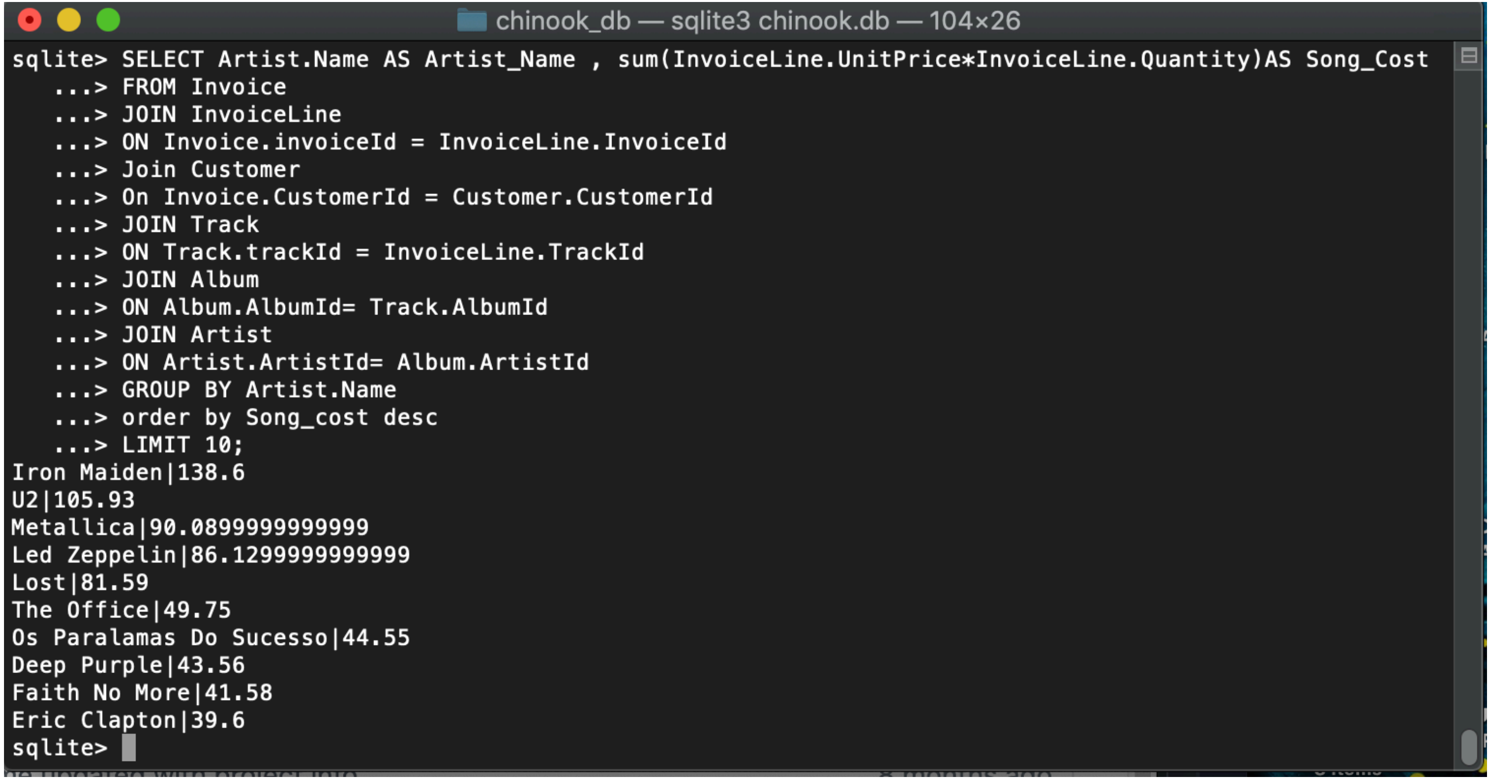
By: Jerome Pullen Jr.

Here are all of the tables within the chinook.db database:



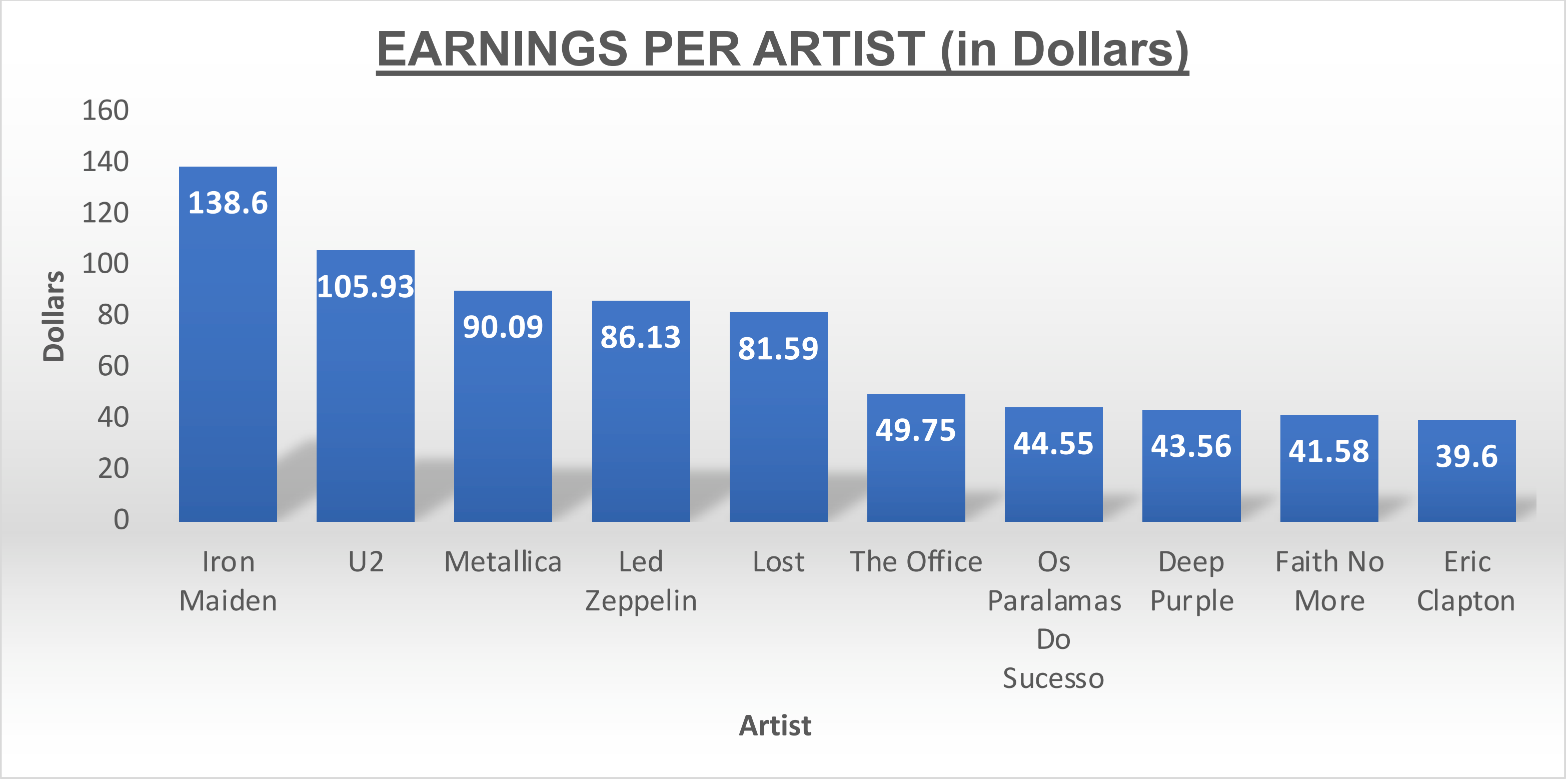
```
sqlite> .tables
Album           Employee        InvoiceLine     PlaylistTrack
Artist          Genre           MediaType       Track
Customer        Invoice         Playlist
sqlite> 
```

I then queried to find which Artists earned the most money:

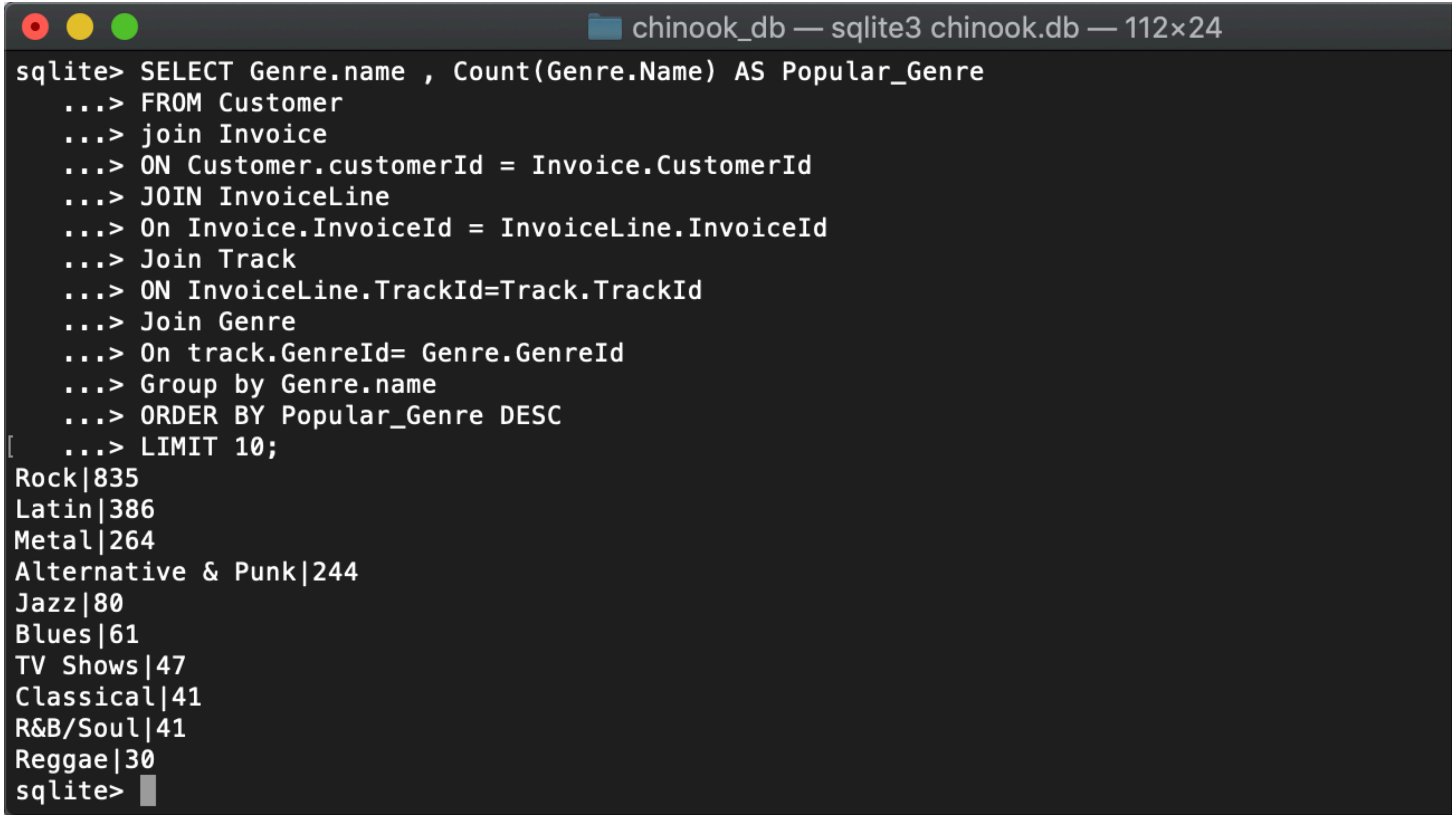


```
sqlite> SELECT Artist.Name AS Artist_Name , sum(InvoiceLine.UnitPrice*InvoiceLine.Quantity)AS Song_Cost
   ...> FROM Invoice
   ...> JOIN InvoiceLine
   ...> ON Invoice.invoiceId = InvoiceLine.InvoiceId
   ...> Join Customer
   ...> On Invoice.CustomerId = Customer.CustomerId
   ...> JOIN Track
   ...> ON Track.trackId = InvoiceLine.TrackId
   ...> JOIN Album
   ...> ON Album.AlbumId= Track.AlbumId
   ...> JOIN Artist
   ...> ON Artist.ArtistId= Album.ArtistId
   ...> GROUP BY Artist.Name
   ...> order by Song_cost desc
   ...> LIMIT 10;
Iron Maiden|138.6
U2|105.93
Metallica|90.0899999999999
Led Zeppelin|86.1299999999999
Lost|81.59
The Office|49.75
Os Paralamas Do Sucesso|44.55
Deep Purple|43.56
Faith No More|41.58
Eric Clapton|39.6
sqlite>
```

ARTISTS WHO EARNED THE MOST MONEY

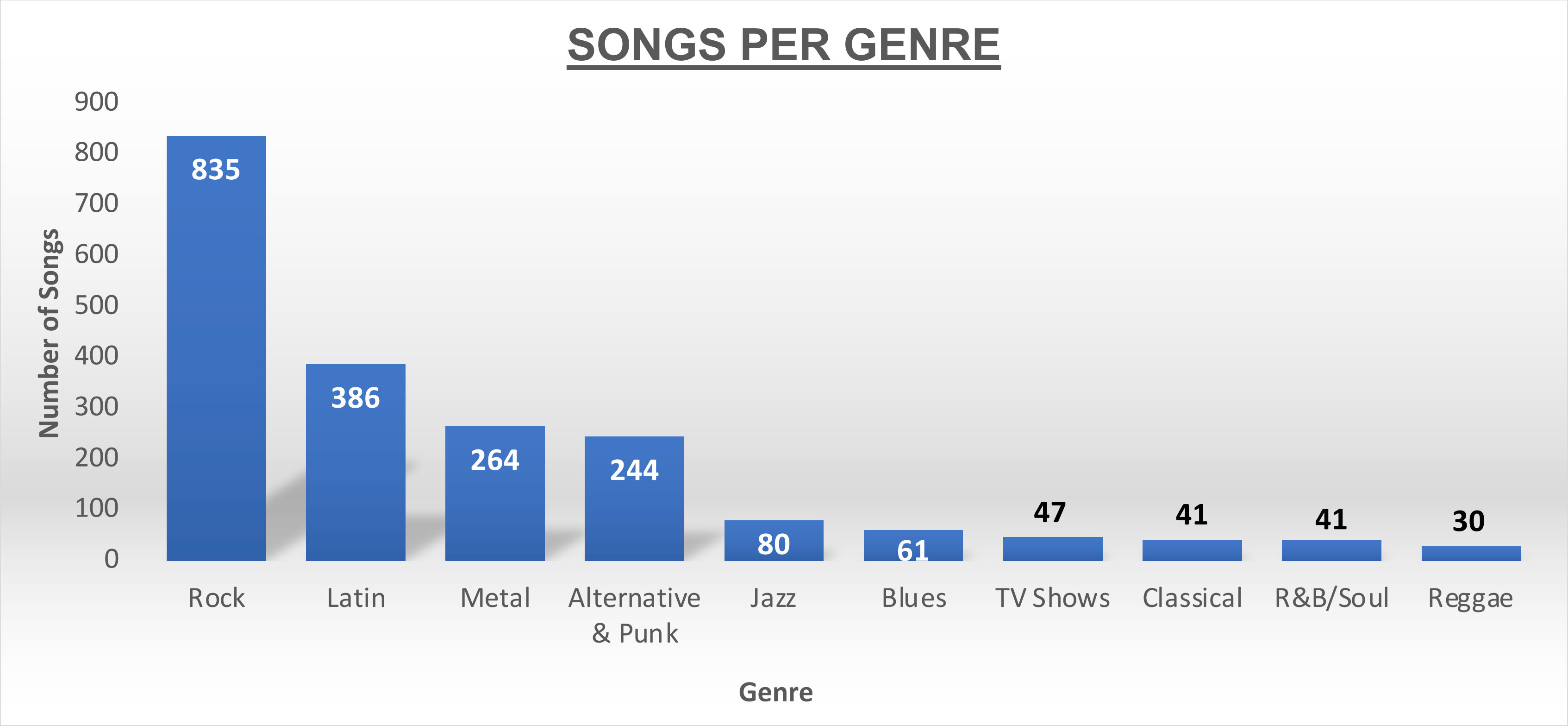


I then queried to find which genres had the most songs:

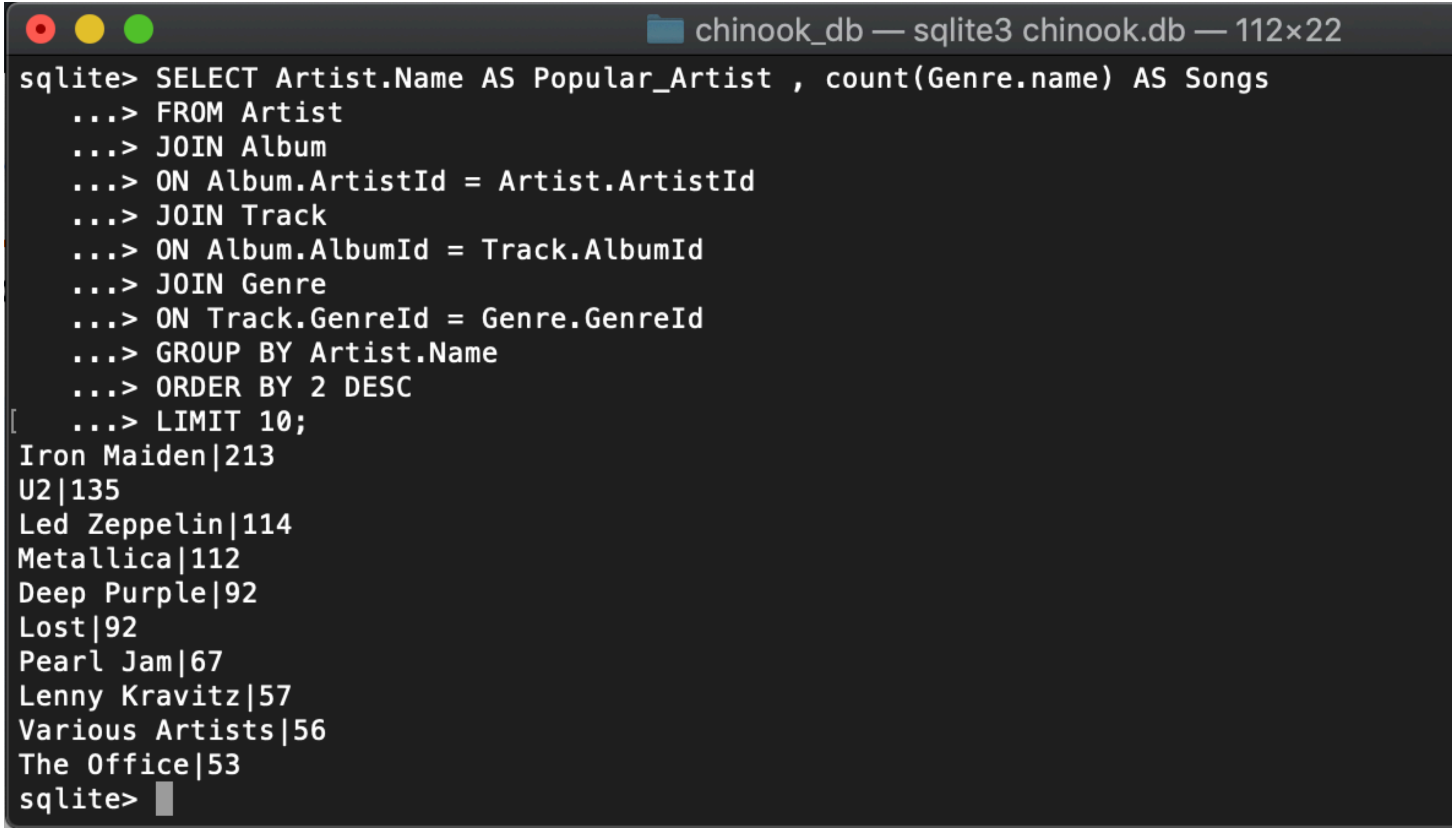


```
sqlite> SELECT Genre.name , Count(Genre.Name) AS Popular_Genre
   ...> FROM Customer
   ...> join Invoice
   ...> ON Customer.customerId = Invoice.CustomerId
   ...> JOIN InvoiceLine
   ...> On Invoice.InvoiceId = InvoiceLine.InvoiceId
   ...> Join Track
   ...> ON InvoiceLine.TrackId=Track.TrackId
   ...> Join Genre
   ...> On track.GenreId= Genre.GenreId
   ...> Group by Genre.name
   ...> ORDER BY Popular_Genre DESC
   ...> LIMIT 10;
Rock|835
Latin|386
Metal|264
Alternative & Punk|244
Jazz|80
Blues|61
TV Shows|47
Classical|41
R&B/Soul|41
Reggae|30
sqlite> 
```

GENRES WITH THE MOST SONGS

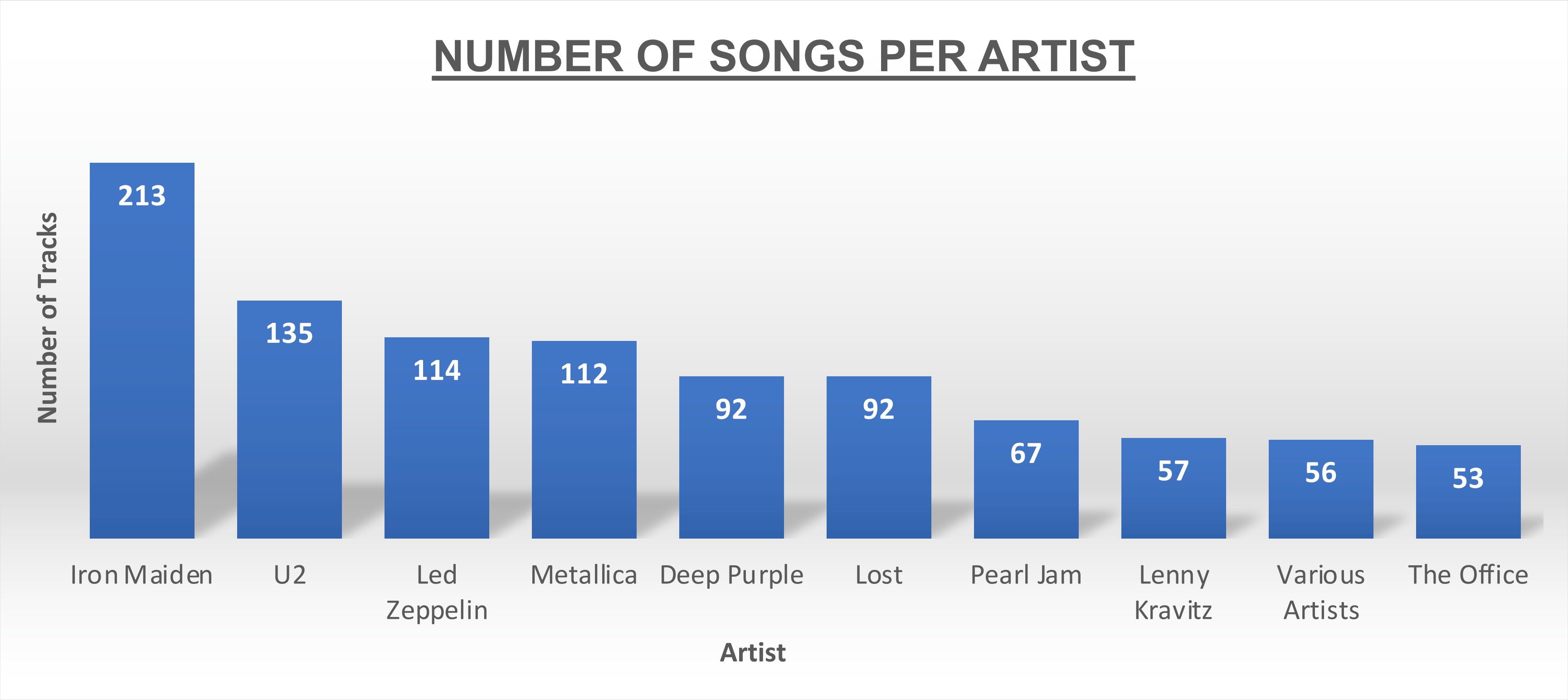


Lastly, I queried to find the Artists with the most songs:



```
sqlite> SELECT Artist.Name AS Popular_Artist , count(Genre.name) AS Songs
   ...> FROM Artist
   ...> JOIN Album
   ...> ON Album.ArtistId = Artist.ArtistId
   ...> JOIN Track
   ...> ON Album.AlbumId = Track.AlbumId
   ...> JOIN Genre
   ...> ON Track.GenreId = Genre.GenreId
   ...> GROUP BY Artist.Name
   ...> ORDER BY 2 DESC
   ...> LIMIT 10;
Iron Maiden|213
U2|135
Led Zeppelin|114
Metallica|112
Deep Purple|92
Lost|92
Pearl Jam|67
Lenny Kravitz|57
Various Artists|56
The Office|53
sqlite>
```

ARTISTS WITH THE MOST SONGS


NUMBER OF SONGS PER ARTIST
Number of Tracks
Iron Maiden
U2
Led Zeppelin
Metallica
Deep Purple
Lost
Pearl Jam
Lenny Kravitz
Various Artists
The Office
Artist