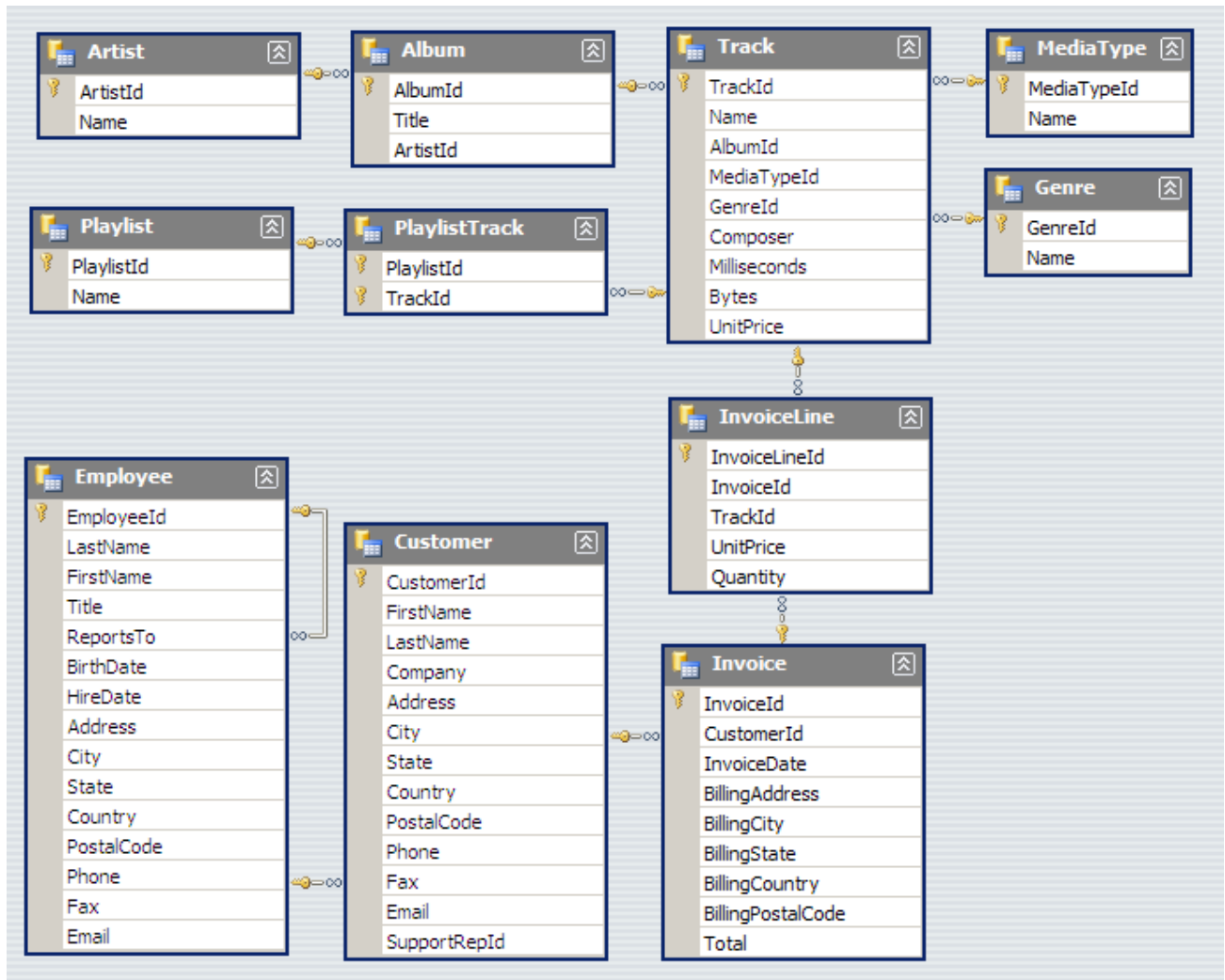


## Chinook Music Database schema:



SQL Queries are written on DB Browser (SQLite syntax)

Visualizations created in Jupyter Notebook with matplotlib library of Python.

**Who are the top 10 artists contributing the most to total revenue, and how much has each contributed?**

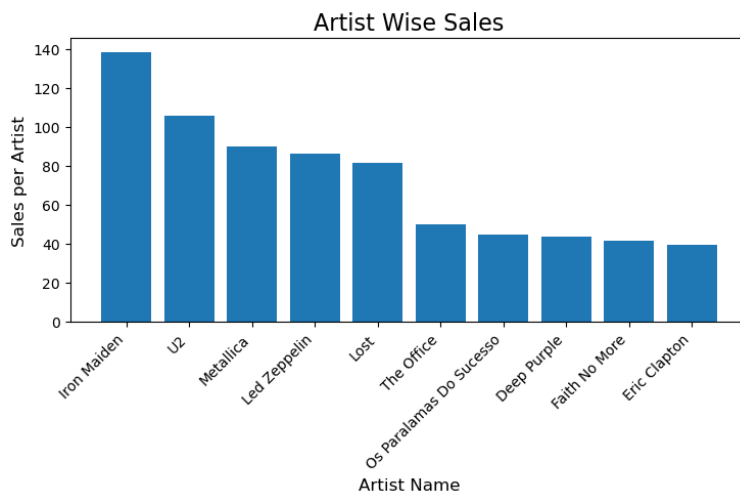
Query:

```
1  /* Which artists make up the top 10 contributors to the total revenue, and how much has each one contributed? */
2
3  SELECT Artist.Name AS artist_name, SUM(InvoiceLine.UnitPrice*InvoiceLine.Quantity) sales_per_artist
4  FROM Track
5  JOIN Album ON Album.AlbumId = Track.AlbumId
6  JOIN Artist ON Artist.ArtistId = Album.ArtistId
7  JOIN InvoiceLine ON InvoiceLine.TrackId = Track.TrackId
8  JOIN Invoice ON Invoice.InvoiceId = InvoiceLine.InvoiceId
9  GROUP BY Artist.Name
10 ORDER BY sales_per_artist DESC
11 LIMIT 10;
```

## Output:

	artist_name	sales_per_artist
1	Iron Maiden	138.6
2	U2	105.93
3	Metallica	90.09
4	Led Zeppelin	86.13
5	Lost	81.59
6	The Office	49.75
7	Os Paralamas Do Sucesso	44.55
8	Deep Purple	43.56
9	Faith No More	41.58
10	Eric Clapton	39.6

## Visualization:



## Summary:

The band Iron **Maiden** **leads** the top contributor chart at Chinook Music Store, closely followed by **U2** and **Lost**.

Across the bar graph, we see a steady decline in contributions with no steep fluctuations across the artists. This data indicates a close competition among the artists.

If Chinook plans to host an event, this will give them an insight into how to allocate their funds while inviting artists to perform.

## Which albums hold the top 10 most sold songs?

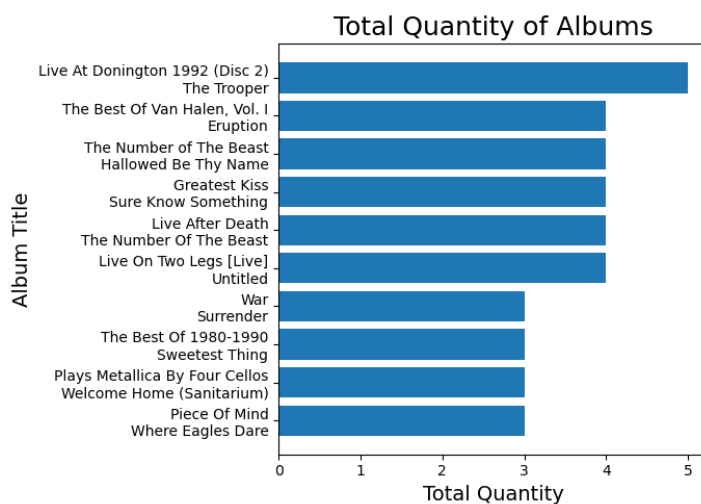
### Query:

```
17 SELECT Album.Title AS album_title, Track.Name AS song_name, Artist.Name AS artist_name, SUM(InvoiceLine.Quantity) AS total_qty
18 FROM Track
19 JOIN Album ON Album.AlbumId = Track.AlbumId
20 JOIN Artist ON Artist.ArtistId = Album.ArtistId
21 JOIN InvoiceLine ON InvoiceLine.TrackId = Track.TrackId
22 JOIN Invoice ON Invoice.InvoiceId = InvoiceLine.InvoiceId
23 GROUP BY Track.Name
24 ORDER BY total_qty DESC
25 LIMIT 10;
```

## Output:

	album_title	song_name	artist_name	total_qty
1	Live At Donington 1992 (Disc 2)	The Trooper	Iron Maiden	5
2	Live On Two Legs [Live]	Untitled	Pearl Jam	4
3	Live After Death	The Number Of The Beast	Iron Maiden	4
4	Greatest Kiss	Sure Know Something	Kiss	4
5	The Number of The Beast	Hallowed Be Thy Name	Iron Maiden	4
6	The Best Of Van Halen, Vol. I	Eruption	Van Halen	4
7	Piece Of Mind	Where Eagles Dare	Iron Maiden	3
8	Plays Metallica By Four Cellos	Welcome Home (Sanitarium)	Apocalyptica	3
9	The Best Of 1980-1990	Sweetest Thing	U2	3
10	War	Surrender	U2	3

## Visualization:



## Summary:

The artist band **Iron Maiden** leads as the top contributor to revenue at Chinook Music Store, closely followed by **U2** and **Lost**.

Across the bar graph, we see a steady decline in contributions with no steep fluctuations across the artists. This indicates a close competition among the artists.

If Chinook plans to host an event, this will give them insight into how to allocate their funds while inviting artists to perform.

## Which are the top 10 genres sold (by quantity)?

### Query:

```

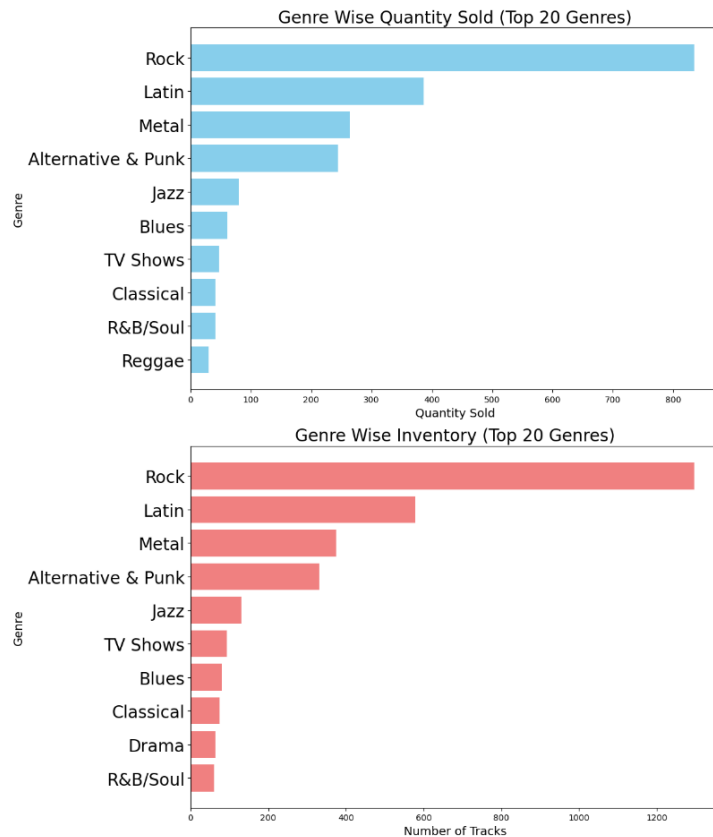
31 WITH GenreSales AS (
32     SELECT
33         g.Name AS genre,
34         SUM(il.Quantity) AS qty_sold
35     FROM Genre g
36     JOIN Track t ON g.GenreId = t.GenreId
37     JOIN InvoiceLine il ON t.TrackId = il.TrackId
38     GROUP BY g.Name
39 ),
40 GenreSongs AS (
41     SELECT
42         g.Name AS genre,
43         COUNT(t.TrackId) AS songs_in_genre
44     FROM Genre g
45     JOIN Track t ON g.GenreId = t.GenreId
46     GROUP BY g.Name
47 )
48 SELECT
49     gs.genre,
50     gs.qty_sold,
51     gs2.songs_in_genre
52 FROM GenreSales gs
53 JOIN GenreSongs gs2 ON gs.genre = gs2.genre
54 ORDER BY gs.qty_sold DESC
55 LIMIT 15;
56

```

## Output:

	genre	qty_sold	songs_in_genre
1	Rock	835	1297
2	Latin	386	579
3	Metal	264	374
4	Alternative & Punk	244	332
5	Jazz	80	130
6	Blues	61	81
7	TV Shows	47	93
8	Classical	41	74
9	R&B/Soul	41	61
10	Reggae	30	58
11	Drama	29	64
12	Pop	28	48
13	Sci Fi & Fantasy	20	26
14	Soundtrack	20	43
15	Hip Hop/Rap	17	35

## Visualization:



### Summary:

By grouping the sales according to the genre reveals that **Rock** has sold an overwhelmingly high number of copies.

However, the data further shows that the sales are proportional to the number of tracks in each genre's collection.

This indicates that the disparity in sales among different genres is primarily due to the greater availability of tracks.

## Which artists have the highest number of albums?

### Query:

```

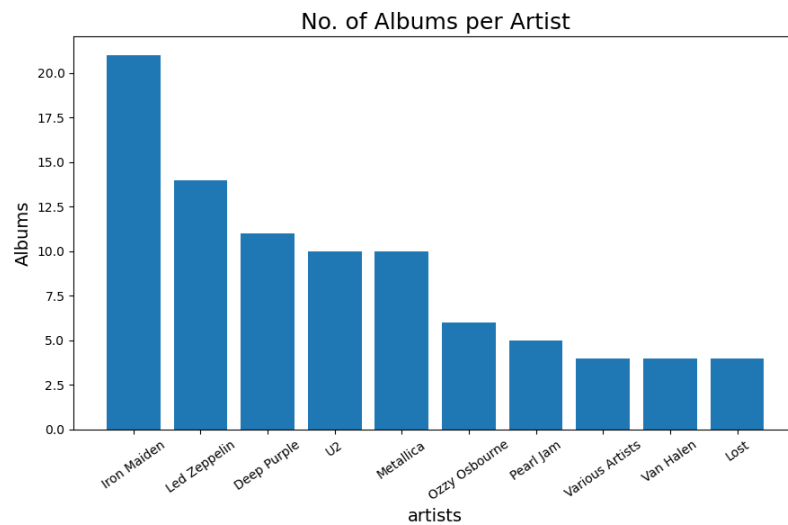
61 SELECT Artist.Name AS artist_name, COUNT(Album.AlbumId) AS album_count
62 FROM Album
63 JOIN Artist ON Artist.ArtistId = Album.ArtistId
64 GROUP BY Artist.Name
65 ORDER BY album_count DESC
66 LIMIT 10;
67

```

### Output:

	artist_name	album_count
1	Iron Maiden	21
2	Led Zeppelin	14
3	Deep Purple	11
4	U2	10
5	Metallica	10
6	Ozzy Osbourne	6
7	Pearl Jam	5
8	Various Artists	4
9	Van Halen	4
10	Lost	4

### Visualization:



### Summary:

**Iron Maiden** is the artist with the highest number of albums, leading with 21 albums. Following them are **Led Zeppelin** and **Deep Purple** with 14 and 11 albums respectively.

### Which are the top 3 selling artists of each genre?

Query:

```

71 WITH ArtistGenreSales AS (
72     SELECT
73         g.Name AS GenreName,
74         ar.Name AS ArtistName,
75         SUM(il.Quantity) AS TotalQuantitySold
76     FROM Genre g
77     JOIN Track t ON g.GenreId = t.GenreId
78     JOIN Album al ON t.AlbumId = al.AlbumId
79     JOIN Artist ar ON al.ArtistId = ar.ArtistId
80     JOIN InvoiceLine il ON t.TrackId = il.TrackId
81     GROUP BY g.Name, ar.Name
82 ),
83 RankedArtists AS (
84     SELECT
85         GenreName,
86         ArtistName,
87         TotalQuantitySold,
88         RANK() OVER (PARTITION BY GenreName ORDER BY TotalQuantitySold DESC) AS Rank
89     FROM ArtistGenreSales
90 ),
91 TotalSalesPerGenre AS (
92     SELECT
93         GenreName,
94         SUM(TotalQuantitySold) AS GenreTotalSales
95     FROM RankedArtists
96     GROUP BY GenreName
97 ),
98
99 SELECT
100     ra.GenreName,
101     ra.ArtistName,
102     ra.TotalQuantitySold,
103     ra.Rank
104 FROM RankedArtists ra
105 JOIN TotalSalesPerGenre tspg ON ra.GenreName = tspg.GenreName
106 WHERE ra.Rank <= 2
107 ORDER BY tspg.GenreTotalSales DESC, ra.Rank;

```

Output:

	GenreName	ArtistName	TotalQuantitySold	Rank
1	Rock	U2	91	1
2	Rock	Led Zeppelin	87	2
3	Latin	Os Paralamas Do Sucesso	45	1
4	Latin	Chico Buarque	27	2
5	Metal	Metallica	91	1
6	Metal	Iron Maiden	70	2
7	Alternative & Punk	Titãs	34	1
8	Alternative & Punk	Faith No More	33	2
9	Alternative & Punk	Green Day	33	2
10	Jazz	Spyro Gyra	20	1
11	Jazz	Gene Krupa	17	2
12	Jazz	Miles Davis	17	2
13	Blues	Eric Clapton	27	1
14	Blues	The Black Crowes	14	2
15	TV Shows	Lost	23	1
16	TV Shows	The Office	16	2
17	Classical	Michael Tilson Thomas & San ...	3	1
18	R&B/Soul	Amy Winehouse	12	1
19	R&B/Soul	James Brown	12	1
20	Classical	Academy of St. Martin in the Fields...	2	2
21	Classical	Berliner Philharmoniker & Herbert ...	2	2

**Summary:**

We see a correlation between the total sales of the top performers and the rank of each genre as a whole. However, we see a break of pattern for the Latin Genre. The top 2 artists of Metal,

**Metallica** and **Iron Maiden** both have a significantly higher number of quantity sold than Latin artists **Os Paralamas Do Sucesso** and **Chico Buarque**.