

SQL Project Report: Chinook Music Store Analysis

Sql script link here <https://github.com/B612Spac/Chinook-Sql-Analysis>

Background

This project analyzes the Chinook music store database to uncover insights into sales performance, customer behavior, genre popularity, and employee effectiveness. Using **SQL for querying** and **Power BI for visualization**, the analysis highlights four key findings: sales have gradually declined despite temporary peaks, high-value customers are concentrated in the USA and select niche markets, Rock and Latin dominate sales while other genres show growth potential, and top-performing employees consistently manage high-value clients.

Data Cleaning & Validation

Before beginning the analysis, I validated the Chinook database to ensure accuracy and reliability of results. Data quality checks included:

1. Duplicate Checks

- Verified there were no duplicate records in key tables such as Customer and Invoice.

Example SQL:

```
SELECT CustomerId, COUNT(*)  
FROM Customer  
GROUP BY CustomerId  
HAVING COUNT(*) > 1;
```

2. Missing / Null Values

- Checked for nulls in critical fields such as customer names, emails, invoice totals, and billing country.

Example SQL:

```
SELECT *  
FROM Invoice  
WHERE BillingCountry  
IS NULL OR Total IS NULL;
```

3. Foreign Key Validation

- Ensured relational integrity by confirming that all invoices link to valid customers, and all tracks link to valid albums and genres.

Example SQL:

```
SELECT i.InvoiceId
FROM Invoice i
```

```
LEFT JOIN Customer c ON i.CustomerId = c.CustomerId
WHERE c.CustomerId IS NULL;
```

Findings

- No duplicates were detected.
- No critical null values were found.
- All foreign key relationships were intact.

Project Objectives:

Your task is to:

1. Analyze sales performance over time

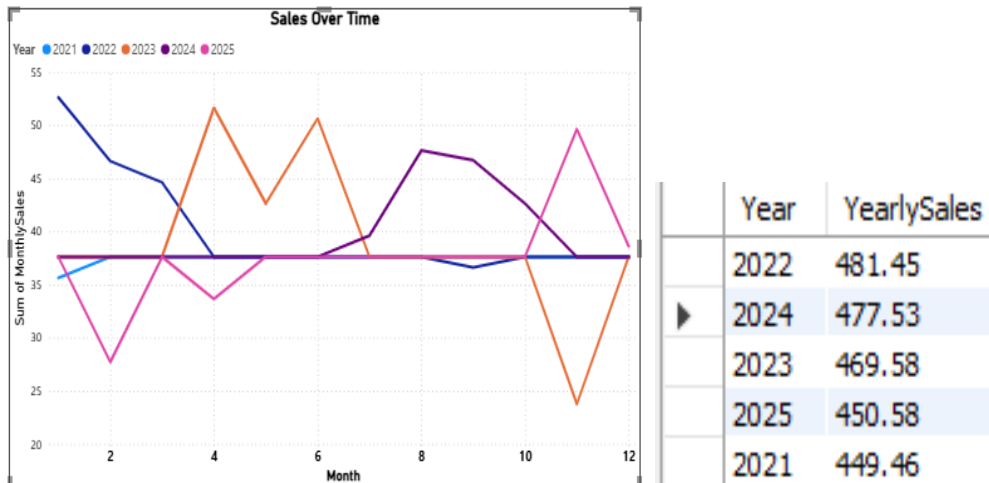
The monthly sales trend over the past 5 years shows that **January sales averaged 37.1 units** across all years **except 2022**. In January 2022, there was a **significant spike** above this average, but sales quickly **returned to the baseline** in subsequent months.

Using this query,

```
SELECT
    YEAR(InvoiceDate) AS Year,
    SUM(Total) AS MonthlySales
FROM Invoice
GROUP BY YEAR(InvoiceDate)
ORDER BY Year;
```

The yearly totals confirm this:

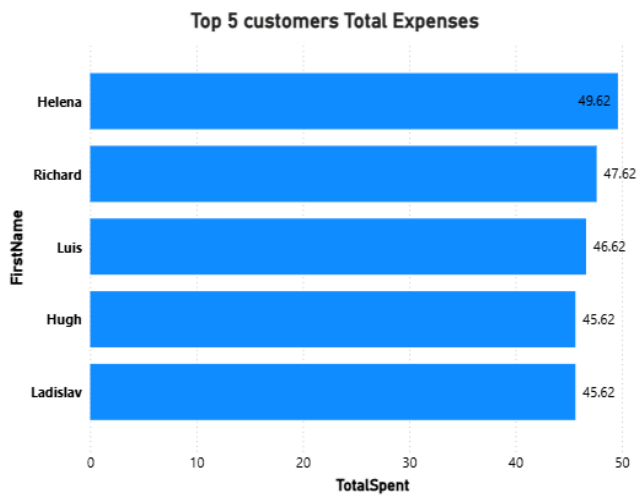
2022 recorded the highest yearly sales at **481.45**, followed by 2024 (**477.53**) and 2023 (**469.58**) with 2021 and 2025 recording the lowest totals at **449.46** and **450.58** respectively. That is to say, although 2022 stands out as a peak year, overall sales performance shows a **gradual decline**. Temporary increases (e.g., in 2022 and 2024) were not sustained and sales consistently reverted to the long-term average.



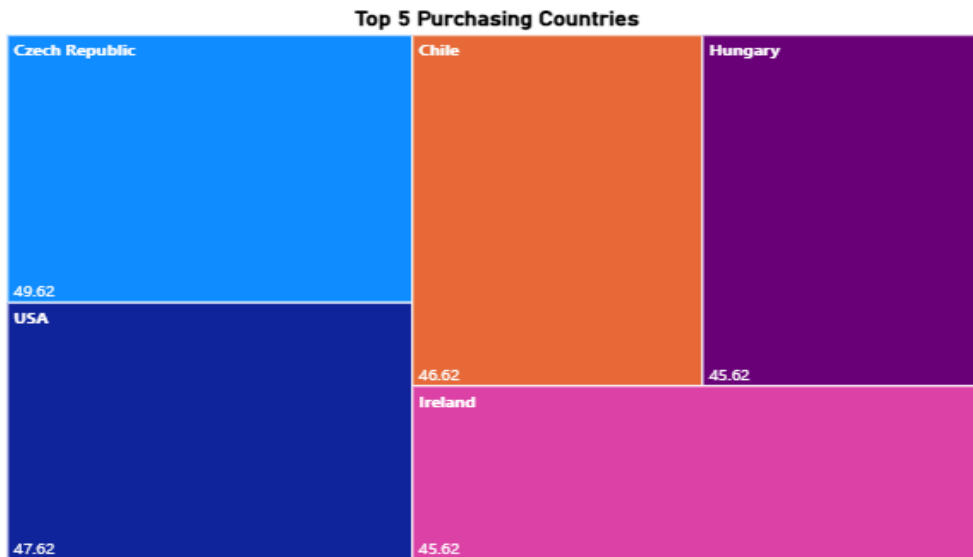
2. Identify high-value customers and their behavior.

Analysis of customer spending shows that high-value customers are concentrated in the **Czech Republic, USA, Chile, Hungary, and Ireland**. Among these, the **USA leads with 13 customers**, contributing the 2nd largest share of sales while the **Czech Republic ranks second** despite having only 2 customers yet is the country with the largest sales. This indicates very high spending per individual. Chile, Hungary, and Ireland also appear in the top 5, but with fewer customers compared to the USA.

This suggests that while the USA has a **broad base of valuable customers**, countries like the Czech Republic represent **small but highly profitable markets**. The difference highlights two customer behaviors: **High-volume contributors** (USA) and **High-intensity spenders** (Czech Republic, with fewer customers but higher average sales)

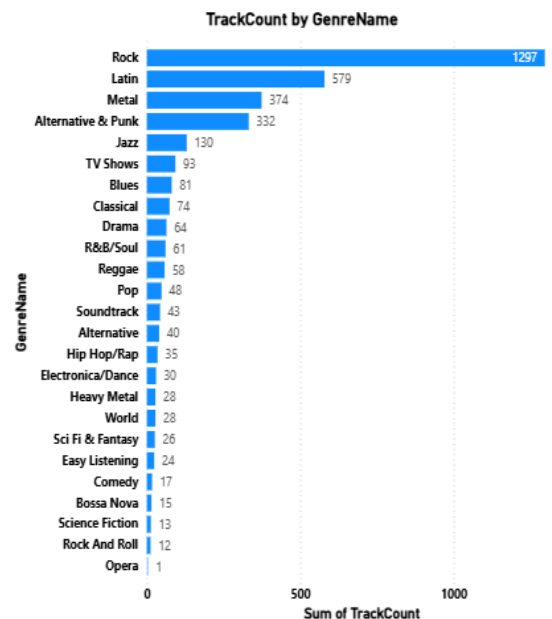
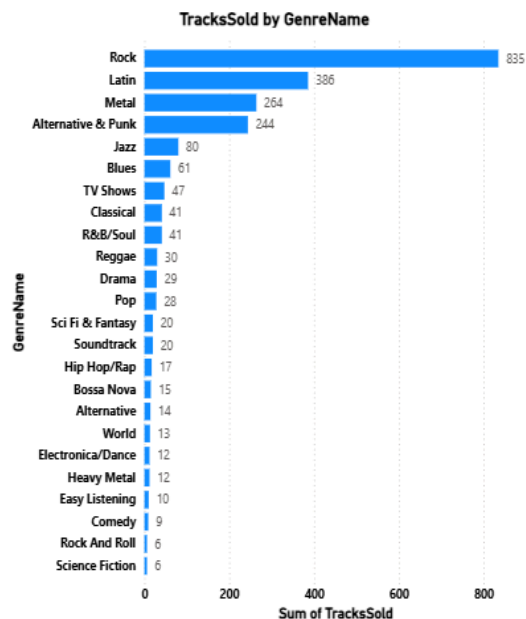


Country	CustomerCount
USA	13
Canada	8
Brazil	5
France	5
Germany	4
United Kingdom	3
Czech Republic	2
Portugal	2
India	2
Norway	1
Austria	1
Belgium	1
Denmark	1
Finland	1
Hungary	1



3. Discover the most popular genres.

From this illustration, I deduced that the top 5 most popular genres are Rock, Latin, Metal, Alternative & Punk, Jazz. There is a strong alignment between inventory levels and sales demand in these genres, suggesting that the store strategically maintains higher stock for faster-moving genres. Rock, in particular, dominates both in catalog size and sales volume, showing it is the main driver of revenue.



4. Evaluate employee performance based on customer support and sales.

	EmployeeId	Emp_Firstname	Emp_Lastname	Cus_Firstname	Cus_Lastname	Revenue
▶	5	Steve	Johnson	Helena	Holý	49.62
	4	Margaret	Park	Richard	Cunningham	47.62
	5	Steve	Johnson	Luis	Rojas	46.62
	3	Jane	Peacock	Ladislav	Kovács	45.62
	3	Jane	Peacock	Hugh	O'Reilly	45.62
	5	Steve	Johnson	Julia	Barnett	43.62
	3	Jane	Peacock	Fynn	Zimmermann	43.62
	3	Jane	Peacock	Frank	Ralston	43.62
	5	Steve	Johnson	Victor	Stevens	42.62
	5	Steve	Johnson	Astrid	Gruber	42.62

The analysis highlights that employees **Steve Johnson, Margaret Park, and Jane Peacock** are associated with the highest-spending customers. Steve Johnson in particular supports five of the top 10 high-revenue customers, while Jane Peacock also manages multiple high-value accounts.

Employees with multiple high-value customers like Steve Johnson and Jane Peacock show strong client management and consistent performance. Employees linked to fewer but high-spending customers like Margaret Park may rely on fewer accounts, which could be a risk if those customers churn.

Recommendations

- **Sales Performance:**
 - a. Investigate the factors behind the 2022 sales spike and replicate them.
 - b. Use seasonal promotions and targeted marketing to reverse the declining sales trend.
- **High-Value Customers:**
 - a. Strengthen retention strategies for small but high-spending markets (e.g., Czech Republic).
 - b. Implement loyalty programs and upselling strategies in the USA to maximize the large customer base.
- **Popular Genres:**
 - a. Continue prioritizing Rock and Latin, the strongest revenue drivers.
 - b. Reallocate resources from low-performing genres to support high-growth potential in Jazz and Alternative & Punk.
- **Employee Performance:**
 - a. Leverage top performers (e.g., Steve Johnson and Jane Peacock) as mentors for other employees.
 - b. Provide training and diversify high-value accounts for employees managing fewer or lower-value clients.

The recommendations focus on replicating past sales drivers, strengthening retention and upselling strategies, reallocating resources to profitable genres, and leveraging top employees as mentors. Together, these strategies provide actionable steps to drive revenue growth, improve customer relationships, and optimize team performance.

Conclusion

This project used SQL and Power BI to analyze the Chinook database, uncovering trends in sales, customer behavior, genre performance, and employee contribution. The findings show declining sales, high-value customers concentrated in specific markets, and the dominance of Rock and Latin genres. Actionable recommendations were made to improve revenue, strengthen customer engagement, and optimize employee performance. This project demonstrates my ability to use data analysis to turn business questions into clear, practical insights.