

# Apple iTunes Music Analysis Through MySQL

## 1. DATA LOADING:

### Step 1: Create Tables

Tables were created for each entity (Artist, Album, Genre, Media\_type, Playlist, Playlist\_track, Employee, Customer, Invoice, Invoice\_line) with appropriate primary keys, foreign keys, and data types.

### Step 2: Load Data

Data was loaded using LOAD DATA INFILE from .csv files stored in the local file system. Special care was taken using:

- FIELDS TERMINATED BY ','
- LINES TERMINATED BY '\r\n'
- NULLIF(..., "") for handling blank fields
- STR\_TO\_DATE for date parsing

## 2. DATA PREPROCESSING:

### 2.1 Data Type Verification

Used DESCRIBE for each table to confirm that all columns have correct data types such as INT, VARCHAR, DATE, and FLOAT.

### 2.2 Missing Value Checks

Used SUM(column IS NULL) on all major fields across all tables.

### 2.3 Missing Value Treatment

- **Customer table:** Filled missing Company, State, Phone, Fax, Postal Code with placeholders like 'Unknown' or 'N/A'.
- **Employee table:** Removed rows with reports\_to IS NULL as they indicate incomplete hierarchy.

## 2.4 Duplicate Handling

Ran grouped COUNT(\*) queries and deleted duplicates using subqueries:

- Playlist and Invoice\_line tables had duplicates removed using NOT IN (SELECT MIN(...)).

## DATASET INFORMATION:

### ◆ 1. Artist

Column Name	Data Type	Description
artist_id	INT (PK)	Unique ID for each artist
name	VARCHAR	Name of the artist

Represents musicians or bands that have created albums.

### ◆ 2. Album

Column Name	Data Type	Description
album_id	INT (PK)	Unique ID for each album
title	VARCHAR	Album title
artist_id	INT (FK)	Links to the artist who created the album

Each album belongs to one artist.

◆ 3. Genre

Column Name	Data Type	Description
genre_id	INT (PK)	Unique ID for each music genre
name	VARCHAR	Genre name (e.g., Rock, Jazz, Pop)

Represents the category or style of music.

◆ 4. Media\_type

Column Name	Data Type	Description
media_type_id	INT (PK)	Format type (e.g., MPEG audio, AAC audio)
name	VARCHAR	Media type name

Describes the encoding or format of music files.

◆ 5. Playlist

Column Name	Data Type	Description
playlist_id	INT (PK)	Unique ID for each playlist

<b>name</b>	VARCHAR	Playlist name (e.g., "My Favorites")
-------------	---------	--------------------------------------

User-defined collections of tracks.

◆ 6. Playlist\_track

Column Name	Data Type	Description
<b>playlist_id</b>	INT (PK, FK)	References a playlist
<b>track_id</b>	INT (PK)	ID of the track in the playlist

Links tracks to playlists (many-to-many relationship).

◆ 7. Employee

Column Name	Data Type	Description
<b>employee_id</b>	INT (PK)	Unique ID for each employee
<b>first_name</b>	VARCHAR	Employee's first name
<b>last_name</b>	VARCHAR	Employee's last name
<b>title</b>	VARCHAR	Job title (e.g., Sales Manager)

<code>reports_to</code>	INT (FK)	Manager (self-referencing FK)
<code>birthdate</code>	DATE	Date of birth
<code>hire_date</code>	DATE	Date of hiring
<code>address to email</code>	VARCHAR	Contact details

Represents employees of the music store, including sales reps and managers.

- ◆ 8. Customer

Column Name	Data Type	Description
<code>customer_id</code>	INT (PK)	Unique customer ID
<code>first_name</code>	VARCHAR	First name
<code>last_name</code>	VARCHAR	Last name
<code>company</code>	VARCHAR	Company name (if any)
<code>address to email</code>	VARCHAR	Contact details
<code>support_rep_id</code>	INT (FK)	Employee assigned to this customer

**Represents customers who purchase music from the store.**

◆ **9. Invoice**

Column Name	Data Type	Description
invoice_id	INT (PK)	Unique invoice ID
customer_id	INT (FK)	Customer who made the purchase
invoice_date	DATE	Date of the invoice
billing_address to billing_postal_code	VARCHAR	Billing info
total	FLOAT	Total amount of the invoice

**Contains customer purchase transaction records.**

◆ **10. Invoice\_line**

Column Name	Data Type	Description
invoice_line_id	INT (PK)	Line item ID in an invoice
invoice_id	INT (FK)	Invoice it belongs to

<b>track_id</b>	INT	Track being purchased
<b>unit_price</b>	FLOAT	Price per unit
<b>quantity</b>	INT	Quantity purchased

Breaks down each invoice into items purchased (each track).

### 3. ANALYSIS & INSIGHTS:

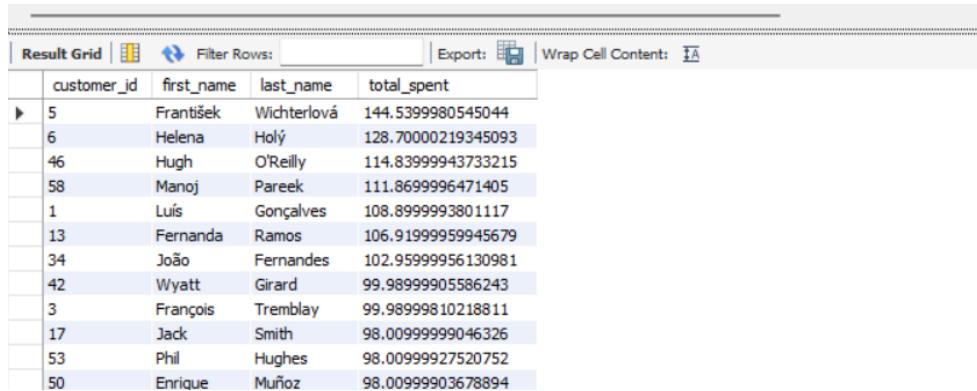
#### 3.1 CUSTOMER ANALYTICS

##### a. Customers who spent the most:

```

7      -- a. Customers who spent the most
8 •  SELECT c.customer_id, c.first_name, c.last_name, SUM(i.total) AS total_spent
9      FROM Customer c
10     JOIN Invoice i ON c.customer_id = i.customer_id
11    GROUP BY c.customer_id
12   ORDER BY total_spent DESC;
13

```



The screenshot shows a database query results grid. The grid has four columns: customer\_id, first\_name, last\_name, and total\_spent. The data is sorted by total\_spent in descending order. The top row shows the column headers. The data starts with customer\_id 5 (František Wichterlová) with a total spent of 144.5399980545044, followed by customer\_id 6 (Helena Holý) with 128.70000219345093, and so on.

	customer_id	first_name	last_name	total_spent
▶	5	František	Wichterlová	144.5399980545044
	6	Helena	Holý	128.70000219345093
	46	Hugh	O'Reilly	114.83999943733215
	58	Manoj	Pareek	111.8699996471405
	1	Luís	Gonçalves	108.8999993801117
	13	Fernanda	Ramos	106.91999959945679
	34	João	Fernandes	102.95999956130981
	42	Wyatt	Girard	99.98999905586243
	3	François	Tremblay	99.98999810218811
	17	Jack	Smith	98.0099999046326
	53	Phil	Hughes	98.00999927520752
	50	Enrique	Muñoz	98.00999903678894

##### Insight:

This SQL query identifies customers who have spent the most, with "František Wichterlová" being the top spender at \$144.54, followed by "Helena Holý" at \$128.70.

## b. Average Customer Lifetime Value:

```
--  
12  -- b. Average Customer Lifetime Value  
13 • SELECT ROUND(AVG(customer_total), 2) AS avg_customer_lifetime_value  
14  FROM (  
15      SELECT c.customer_id, SUM(i.total) AS customer_total  
16      FROM Customer c  
17      JOIN Invoice i ON c.customer_id = i.customer_id  
18      GROUP BY c.customer_id  
19  ) AS customer_values;  
20  
21  -- c. Repeat vs One-time Customers  
22 • SELECT
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	avg_customer_lifetime_value			
▶	79.82			

### Insight:

The average customer lifetime value, calculated by summing individual customer totals and then averaging them, is \$79.82.

## c. Repeat Customers:

```
--  
21  -- c. Repeat vs One-time Customers  
22 • SELECT  
23      CASE WHEN invoice_count > 1 THEN 'Repeat' ELSE 'One-time' END AS customer_type,  
24      COUNT(*) AS total_customers  
25  FROM (  
26      SELECT customer_id, COUNT(*) AS invoice_count  
27      FROM Invoice  
28      GROUP BY customer_id  
29  ) AS counts  
30  GROUP BY customer_type;  
31  
32  -- d. Revenue per Customer per Country  
33 • SELECT country, ROUND(SUM(i.total)/COUNT(DISTINCT c.customer_id), 2) AS revenue_per_customer  
34  FROM Customer c
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	customer_type	total_customers		
▶	Repeat	59		

### Insight:

It shows that there are 59 repeat customers, identified by having more than one invoice.

#### d. Revenue per Customer per Country:

```
34  -- d. Revenue per Customer per Country
35 •   SELECT country, ROUND(SUM(i.total)/COUNT(DISTINCT c.customer_id), 2) AS revenue_per_customer
36   FROM Customer c
37   JOIN Invoice i ON c.customer_id = i.customer_id
38   GROUP BY country
39   ORDER BY revenue_per_customer DESC;
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	country	revenue_per_customer		
▶	Czech Republic	136.62		
	Ireland	114.84		
	Spain	98.01		
	Chile	97.02		
	Portugal	92.56		
	India	91.57		
	Brazil	85.54		
	Germany	83.65		
	United Kingdom	81.84		
	Australia	81.18		
	USA	80.04		
	Finland	79.2		
	Hungary	78.21		
	France	77.81		

#### Insight:

This query results show the "Revenue per Customer per Country" and indicate that the Czech Republic has the highest revenue per customer at \$136.62, followed by Ireland at \$114.84.

#### e. Customers inactive in last 6 months:

```
41  -- e. Customers inactive in last 6 months
42 •   SELECT c.customer_id, c.first_name, c.last_name, MAX(i.invoice_date) AS last_purchase
43   FROM Customer c
44   JOIN Invoice i ON c.customer_id = i.customer_id
45   GROUP BY c.customer_id
46   HAVING MAX(i.invoice_date) < CURDATE() - INTERVAL 6 MONTH
47   order by last_purchase desc;
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	customer_id	first_name	last_name	last_purchase
▶	53	Phil	Hughes	2020-12-30
	20	Dan	Miller	2020-12-29
	33	Ellie	Sullivan	2020-12-27
	52	Emma	Jones	2020-12-27
	55	Mark	Taylor	2020-12-21
	14	Mark	Philips	2020-12-20
	49	Stanislaw	Wójcik	2020-12-20
	59	Rishabh	Mishra	2020-12-19
	15	Jennifer	Peterson	2020-12-18
	35	Madalena	Sampaio	2020-12-17
	46	Hugh	O'Reilly	2020-12-12
	26	Richard	Cunningh...	2020-12-10
	21	Kathy	Chase	2020-12-09
	29	Robert	Brown	2020-12-08
	44	Terhi	Hämäläinen	2020-12-08
	13	Fernanda	Ramos	2020-11-28
	42	Wyatt	Girard	2020-11-27
	23	John	Gordon	2020-11-27
	31	Martha	Silk	2020-11-24
	12	Roberto	Almeida	2020-11-23
	2	Leonie	Köhler	2020-11-21
	16	Frank	Harris	2020-11-20
	51	Joakim	Johansson	2020-11-17

## Insight:

This list presents 59 customers, along with their last purchase dates, all of which fall in 2019 or 2020, with Phil having last purchase on 2020-12-30.

## 3.2 SALES & REVENUE ANALYSIS

### a. Monthly Revenue (Last 2 Years):

```
52      -- a. Monthly Revenue (Last 2 Years)
53 •   SELECT DATE_FORMAT(invoice_date, '%Y-%m') AS month, SUM(total) AS revenue
54     FROM Invoice
55     WHERE invoice_date >= '2019-01-01'
56     GROUP BY month
57     ORDER BY month;
```

Result Grid		
	month	revenue
▶	2019-01	85.13999891281128
	2019-02	59.399999380111694
	2019-03	125.72999978065491
	2019-04	116.81999850273132
	2019-05	104.93999814987183
	2019-06	92.06999921798706
	2019-07	124.73999905586243
	2019-08	159.38999915122986
	2019-09	126.71999907493591
	2019-10	43.55999970436096
	2019-11	71.27999973297119
	2019-12	111.87000036239624
	2020-01	43.55999970436096
	2020-02	97.01999950408936

## Insight:

The query displays monthly revenue starting from January 2019, showing fluctuations with a high of approximately \$159.39 in August 2019 and a low of \$43.56 in October 2019 and January 2020.

### b. Average Invoice Value:

```
59      -- b. Average Invoice Value
60 •   SELECT ROUND(AVG(total), 2) AS avg_invoice_value FROM Invoice;
61
```

Result Grid	
	avg_invoice_value
▶	7.67

### **Insight:**

The average invoice value, rounded to two decimal places, is \$7.67.

### **c. Revenue by Sales Representative contribute:**

```
62 -- c. Revenue by Sales Representative contribute
63 • SELECT e.employee_id, e.first_name, e.last_name, ROUND(SUM(i.total),2) AS total_revenue
64 FROM Employee e
65 JOIN Customer c ON e.employee_id = c.support_rep_id
66 JOIN Invoice i ON c.customer_id = i.customer_id
67 GROUP BY e.employee_id;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

employee_id	first_name	last_name	total_revenue
3	Peacock	Jane	1731.51
4	Park	Margaret	1584
5	Johnson	Steve	1393.92

### **Insight:**

The query results display the "Revenue by Sales Representative," showing that Jane Peacock has generated the highest total revenue of 1731.51, followed by Margaret Park (1584), and Steve Johnson (1393.92).

### **d. Peak Sales Months or Quarters:**

```
69 -- d. Peak Sales Months or Quarters
70 • SELECT
71     DATE_FORMAT(invoice_date, '%Y-%m') AS month,
72     SUM(total) AS revenue
73 FROM Invoice
74 GROUP BY month
75 ORDER BY revenue DESC
76 LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

month	revenue
2018-01	183.14999890327454
2019-08	159.38999915122986
2018-03	148.50000071525574
2020-10	144.53999996185303
2017-04	142.55999970436096

## Insight:

The query identifies the top 5 peak sales months, with January 2018 having the highest revenue at \$183.15, followed by August 2019 at \$159.39.

## 3.3 PRODUCT & CONTENT ANALYSIS

### a. Tracks with Most Revenue:

```
82 -- a. Tracks with Most Revenue
83 • SELECT
84     track_id,
85     SUM(unit_price * quantity) AS revenue
86 FROM
87     Invoice_line
88 GROUP BY
89     track_id
90 ORDER BY
91     revenue DESC
92 LIMIT 10;
93
```

Result Grid | Filter Rows: Export: Wrap Cell Content: □

track_id	revenue
3336	30.690000295639038
1495	13.860000133514404
1489	13.860000133514404
1487	12.870000123977661
6	12.870000123977661
1490	12.870000123977661
1483	11.880000114440918
2563	10.890000104904175
1488	10.890000104904175
1485	10.890000104904175

## Insight:

The query identifies the top 10 tracks by revenue, with track ID 3336 generating the highest revenue of 30.69, significantly more than the subsequent tracks.

### b. Most Purchased Playlists:

```
94 -- b. Most Purchased Playlists
95 • SELECT
96     p.name AS playlist_name,
97     COUNT(il.track_id) AS total_purchases
98 FROM
99     Playlist p
100 JOIN
101     Playlist_track pt ON p.playlist_id = pt.playlist_id
102 JOIN
103     Invoice_line il ON pt.track_id = il.track_id
104 GROUP BY
105     p.playlist_id
106 ORDER BY
107     total_purchases DESC
108 LIMIT 10;
```

Result Grid | Filter Rows: Export: Wrap Cell Content: □

playlist_name	total_purchases
Music	4742
90's Music	1848
Heavy Metal Classic	56
Grunge	44
Classical	43
Classical 101 - Next Steps	16
Classical 101 - The Basics	15
Classical 101 - Deep Cuts	12
Brazilian Music	5
TV Shows	3

### Insight:

The query results reveal that "Music" is by far the most purchased playlist with 4742 purchases, significantly outranking "90's Music" (1848 purchases) and all other playlists, indicating a broad preference for general music compilations.

### c. Tracks Never Purchased:

```
110 -- c. Tracks Never Purchased
111 • SELECT
112     DISTINCT pt.track_id
113     FROM
114         Playlist_track pt
115     LEFT JOIN
116         Invoice_line il ON pt.track_id = il.track_id
117     WHERE
118         il.track_id IS NULL;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows: |

track_id
99
101
104
106
107
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125

### Insight:

The query successfully identifies a list of track IDs, starting from 99 and continuing sequentially, that have never been purchased.

## 3.4 EMPLOYEE & OPERATIONAL EFFICIENCY

### a. Avg Customers per Employee

```
124 -- a. Avg Customers per Employee
125 • SELECT ROUND(AVG(customer_count), 2) AS avg_customers_per_employee
126     FROM (
127         SELECT support_rep_id, COUNT(*) AS customer_count
128             FROM Customer
129             GROUP BY support_rep_id
130     ) AS emp_customers;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows: |

avg_customers_per_employee
19.67

## Insight:

The average number of customers per employee is APPROX 20.

### b. Revenue by Employee Region

```
132      -- b. Revenue by Employee Region (if available via address or city)
133 •  SELECT e.country, SUM(i.total) AS total_revenue
134     FROM Employee e
135     JOIN Customer c ON e.employee_id = c.support_rep_id
136     JOIN Invoice i ON c.customer_id = i.customer_id
137     GROUP BY e.country;
138
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	country	total_revenue		
▶	Canada	4709.429977893829		

## Insight:

The query result indicates that Canada is the only country for which employee-driven revenue data is available, with a total revenue of approximately 4709.43.

## 3.5 GEOGRAPHIC TRENDS

### a. Highest Number of Customers by Country or City

```
143      -- a. Highest Number of Customers by Country or City
144 •  SELECT country, COUNT(*) AS total_customers
145     FROM Customer
146     GROUP BY country
147     ORDER BY total_customers DESC;
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	country	total_customers		
▶	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		
	Ireland	1		
	Italy	1		
	Netherlands	1		
	Poland	1		
	Spain	1		
	Sweden	1		
	Australia	1		
	Argentina	1		
	Chile	1		

### **Insight:**

The query results indicate that the USA has the highest number of customers (13), followed by Canada (8), and then Brazil and France (both with 5 customers each).

### **b. Revenue by Region (Country)**

country	revenue
USA	1040.49
Canada	535.59
Brazil	427.68
France	389.07
Germany	334.62
Czech Republic	273.24
United Kingdom	245.52
Portugal	185.13
India	183.15
Ireland	114.84
Spain	98.01
Chile	97.02
Australia	81.18
Finland	79.2
Hungary	78.21
Poland	76.23
Sweden	75.24
Norway	72.27
Austria	69.3
Netherlands	65.34
Belgium	60.39
Italy	50.49
Argentina	39.6
Denmark	37.62

### **Insight:**

The query results for "Revenue by Region (Country)" show that the USA generates the highest revenue at \$1040.49, followed by Canada with \$535.59 and Brazil with \$427.68.

## **3.6 CUSTOMER RETENTION & PATTERNS**

### **a. Purchase Frequency per Customer**

customer_id	purchase_count
1	13
2	11
3	9
4	9
5	18
6	12
7	9
8	7
9	10
10	12
11	10
12	11
13	15
14	10
15	9
16	8
17	12
18	8
19	9
20	12
21	11
22	12
23	10
24	8

## Insight:

The query results show the purchase frequency for each customer, with customer ID 5 having the highest purchase count of 18, and several customers (e.g., 2, 4, 6) having around 9 to 13 purchases.

### b. Avg Time Between Purchases

```
165 -- b. Avg Time Between Purchases
166 • WITH PurchaseDates AS (
167     SELECT customer_id, invoice_date,
168         LAG(invoice_date) OVER (PARTITION BY customer_id ORDER BY invoice_date) AS prev_purchase
169     FROM Invoice
170 )
171     SELECT customer_id,
172         ROUND(AVG(DATEDIFF(invoice_date, prev_purchase)), 2) AS avg_days_between
173     FROM PurchaseDates
174     WHERE prev_purchase IS NOT NULL
175     GROUP BY customer_id;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

customer_id	avg_days_between
1	106.25
2	129.20
3	147.50
4	136.13
5	74.29
6	104.45
7	133.50
8	131.83
9	120.67
10	112.27
11	108.11
12	137.50
13	92.14
14	128.33
15	166.75
16	185.86
17	104.18
18	165.29

## Insight:

The query calculates the average days between purchases for each customer, showing that customer ID 29 has the longest average interval at 392 days, while customer ID 5 has a much shorter interval at 74.29 (nearly 75) days.

## 3.7 OPERATIONAL OPTIMIZATION

### a. Most Common Track Combinations

```
181 -- a. Most Common Track Combinations
182 • SELECT ill1.track_id AS track1, ill2.track_id AS track2, COUNT(*) AS times_bought_together
183     FROM Invoice_line ill1
184     JOIN Invoice_line ill2 ON ill1.invoice_id = ill2.invoice_id AND ill1.track_id < ill2.track_id
185     GROUP BY track1, track2
186     ORDER BY times_bought_together DESC
187     LIMIT 10;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

track1	track2	times_bought_together
1479	1487	9
1483	1484	9
1485	1495	9
1479	1483	8
1479	1490	8
1479	1484	8
1479	1488	8
1479	1485	8
1479	1480	8
1479	1481	8

### **Insight:**

The query identifies the most common track combinations purchased together, with track IDs (1479, 1487), (1483, 1484), and (1485, 1495) all being bought together 9 times, making them the most frequent pairings.

### **b. Price vs Sales Analysis**

```
189      -- b. Price vs Sales Analysis
190 •   SELECT il.unit_price, COUNT(*) AS total_sales
191     FROM Invoice_line il
192     GROUP BY il.unit_price
193     ORDER BY il.unit_price;
```

A screenshot of a database query results grid. The grid has two columns: 'unit\_price' and 'total\_sales'. There is one row of data with values 0.99 and 4745 respectively. The grid includes standard data manipulation toolbar icons like 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content'.

	unit_price	total_sales
▶	0.99	4745

### **Insight:**

The query reveals that items priced at \$0.99 account for the vast majority of sales, with a total of 4745 units sold at this price point.

## **Conclusion:**

The analysis highlights that top customers like František Wichterlová and Helena Holý significantly impact revenue, while the USA leads in both customer base and revenue generation. Operational efficiencies show that \$0.99-priced items are the most popular, contributing substantially to overall sales. These insights underscore the importance of customer segmentation and pricing strategies in optimizing revenue. Moving forward, focusing on high-value customers and refining product offerings could further enhance profitability and operational effectiveness in the market.

# Apple iTunes Music Analysis Dashboard

## Overview:

This Dashboard analyzes the iTunes dataset using Power BI to extract actionable business insights across revenue, customer behavior, music content, and employee performance. It includes data transformation, Merging , and DAX calculations to build a comprehensive 4-page dashboard.

## Steps Followed:

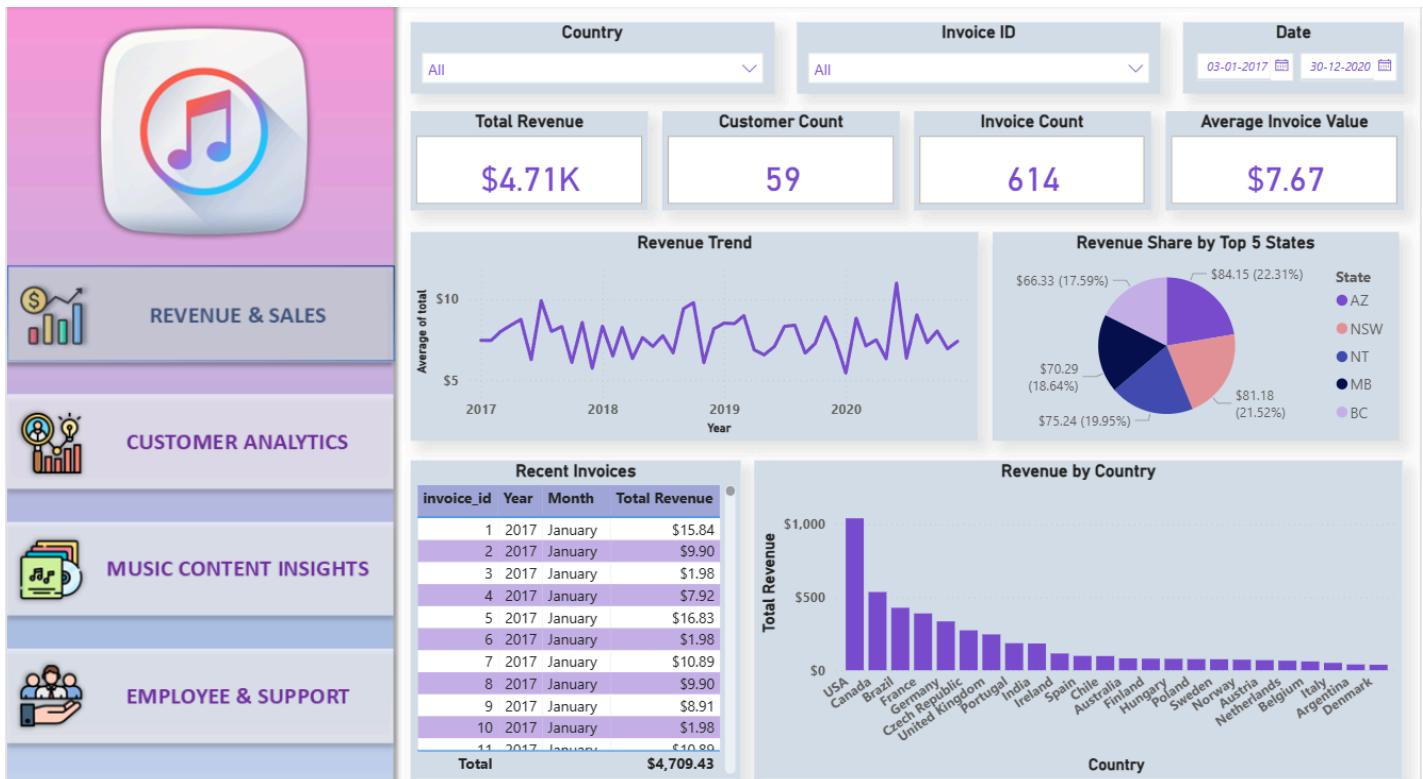
1. Imported multiple clean CSV files into Power BI (e.g., Customer, Invoice, Invoice\_line, Track, Playlist, Employee etc.).
2. Merged the required tables to one table for each page of the dashboard.
3. Added DAX measures for KPIs.
4. Built visualizations based on business objectives.
5. Used slicers and filters for interactivity and segmentation.

## Datasets Used:

Table Name	Description
Customer_cleaned	Customer details including contact, location, and support representative ID.
Employee_cleaned	Company employees, titles, hierarchy ( <code>reports_to</code> ), and hire dates.
Invoice_cleaned	All purchase records including customer ID, invoice date, country/state, and total amount.
Invoice_line_cleaned	Line-level invoice details such as quantity, unit price, and track ID.

Track	Music track information.
Genre_cleaned	Genre ID and name for tracks.
Media_type_cleaned	Media format types for tracks.
Playlist_cleaned	Playlist metadata including names.
Playlist_track_cleaned	Junction table mapping tracks to playlists.
Album_cleaned	Album titles with artist references.
Artist_cleaned	Artist names and IDs.

## PAGE 1 (REVENUE & SALES)



**Dataset used:** Invoice\_cleaned dataset

### **Key Metrics (Key Performance Indicators - KPIs):**

Total Revenue	Customer Count	Invoice Count	Average Invoice Value
\$4.71K	59	614	\$7.67

- **Total Revenue:** \$4.71K (\$4,710)
- **Customer Count:** 59
- **Invoice Count:** 614
- **Average Invoice Value:** \$7.67

### **Filters Available:**

Country	Invoice ID	Date
All	All	03-01-2017  30-12-2020

- **Country Filter:** "All"
- **Invoice ID Filter:** "All"
- **Date Range Filter:** "01-01-2017 - 30-12-2020"

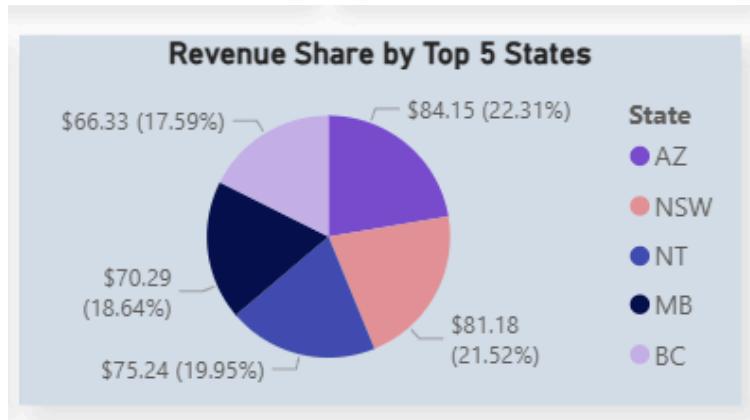
### **Visual Insights:**

#### **1. Revenue Trend (Line Chart)**



- Revenue fluctuated between \$5 and \$10 from 2017 to 2020.
- Noticeable spikes in mid-2019 and again in late 2020 indicate revenue peaks.

## 2. Revenue Share by Top 5 States (Pie Chart)



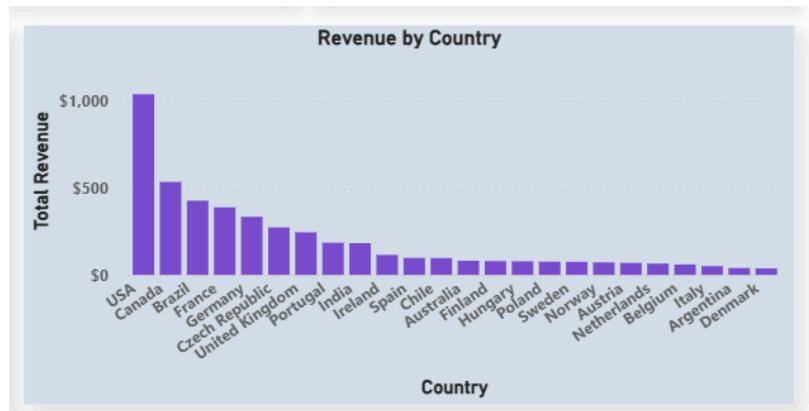
- **Arizona (22.31%)** and **Manitoba (21.52%)** are the leading states in revenue contribution.
- **New South Wales (19.95%), Northwest Territories(18.64%), and British Columbia (17.59%)** follow closely, showing relatively even distribution among top contributors.

## 3. Recent Invoices Table

Recent Invoices				
invoice_id	Year	Month	Total Revenue	
1	2017	January	\$15.84	
2	2017	January	\$9.90	
3	2017	January	\$1.98	
4	2017	January	\$7.92	
5	2017	January	\$16.83	
6	2017	January	\$1.98	
7	2017	January	\$10.89	
8	2017	January	\$9.90	
9	2017	January	\$8.91	
10	2017	January	\$1.98	
11	2017	January	\$10.80	
<b>Total</b>			<b>\$4,709.43</b>	

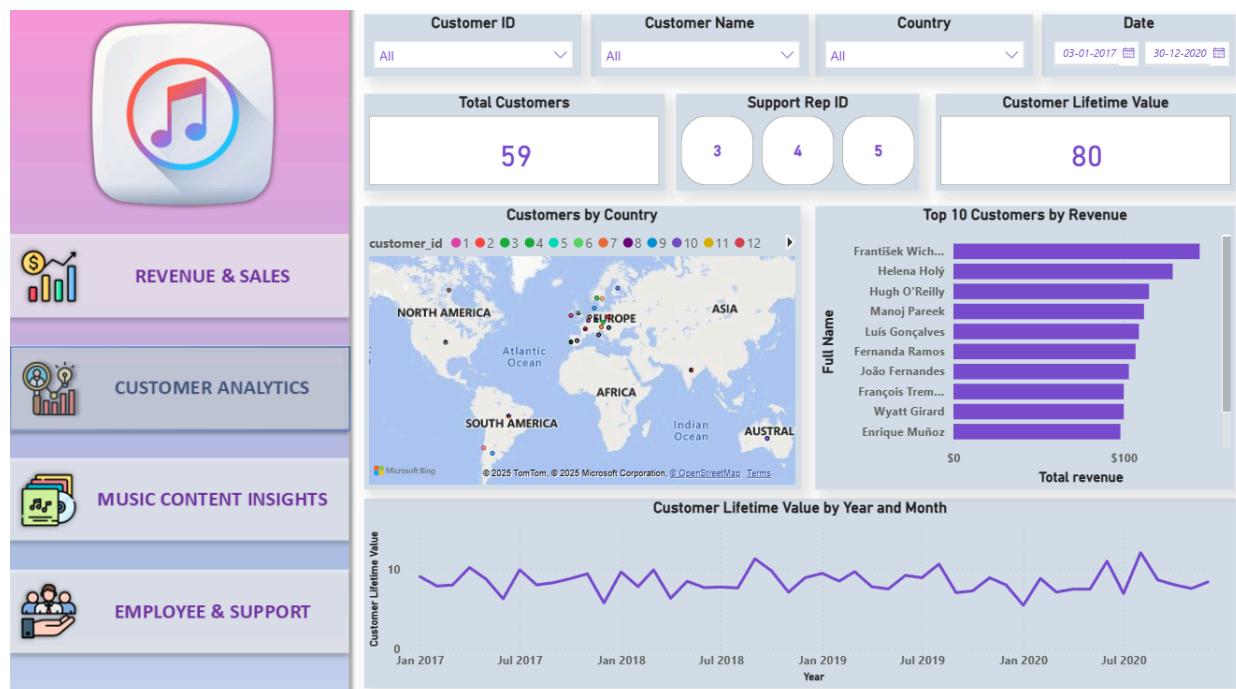
- Displays invoice entries ranging from Jan 2017- Dec 2020.
- Shows values ranging from ~\$0.99 to ~\$23.76.
- The total revenue from these invoices amounts to **\$4,709.43** (validating the KPI total).

#### 4. Revenue by Country (Bar Chart)



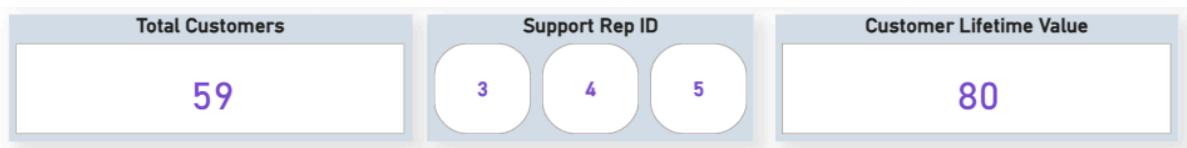
- **USA** generates the highest revenue by far, exceeding \$1,000.
- Other top contributors: **Canada, Brazil, France, and Germany**.
- European countries dominate the mid-tier, while smaller contributions are observed from countries like **Italy, Denmark, and Argentina**.

## PAGE 2 (CUSTOMER ANALYTICS)



**Dataset used:** customer-Invoice\_merged dataset( obtained by merging customer\_cleaned and Invoice\_cleaned dataset )

### **Key Metrics (Key Performance Indicators - KPIs):**



- **Total Customers:** 59
- **Support Rep ID:** 3, 4, 5 (with individual counts or representations)
- **Customer Lifetime Value:** 80

### **Filters Available:**



- **Customer ID Filter:** "All"
- **Customer Name Filter:** "All"
- **Country Filter:** "All"
- **Date Range Filter:** "03-01-2017 - 30-12-2020"

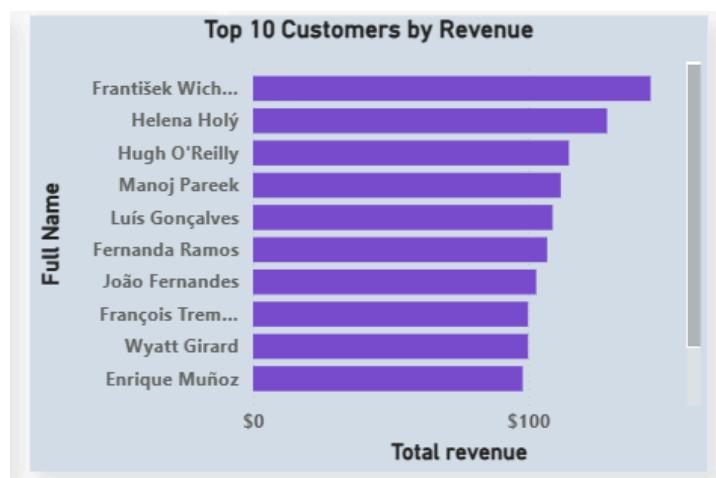
### **Visual Insights:**

#### **1. Customers by Country (Map)**



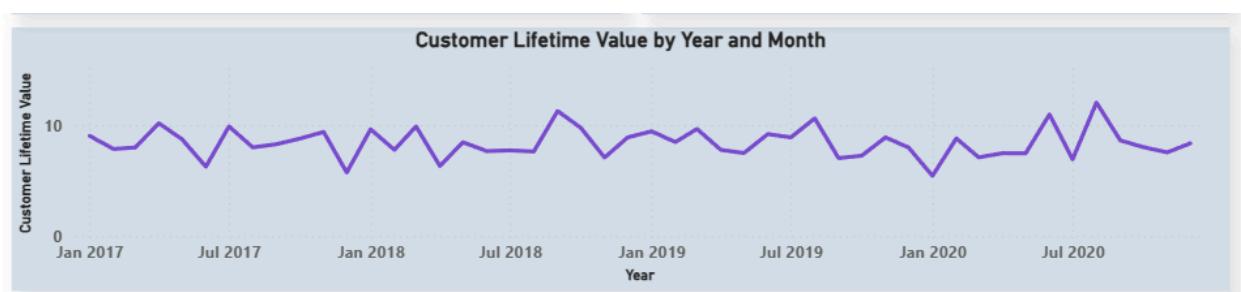
- Customers are distributed globally, with visible clusters in:
  - **Europe** (Germany, France, Czech Republic, etc.)
  - **North America** (USA, Canada)
  - **South America, Asia, and Australia** also have small customer footprints.

## 2. Top 10 Customers by Revenue (Bar Chart)



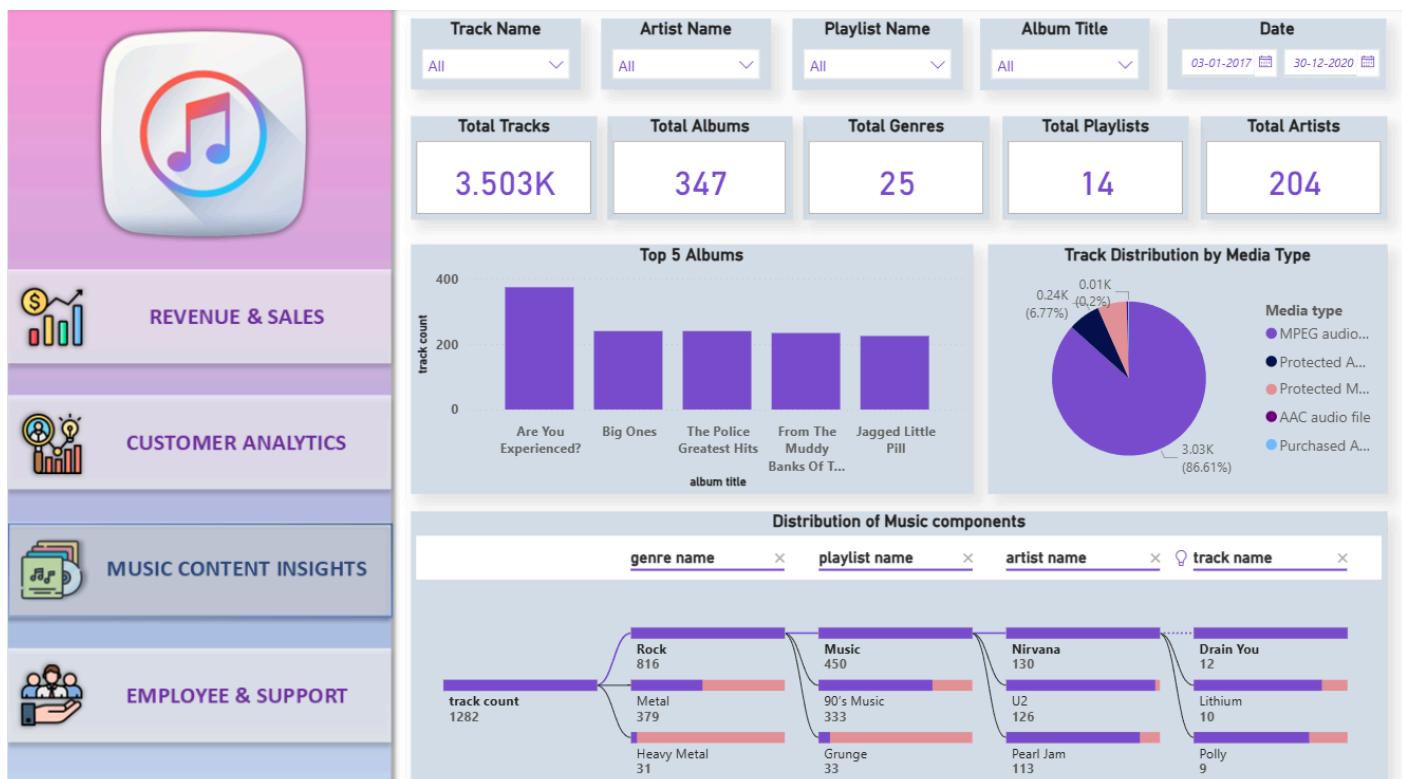
- **František Wichterlova** is the highest-paying customer, exceeding \$100 in total spend.
- Others like **Helena Holý** and **Hugh O'Reilly** follow.

## 3. Customer Lifetime Value Over Time (Line Chart)



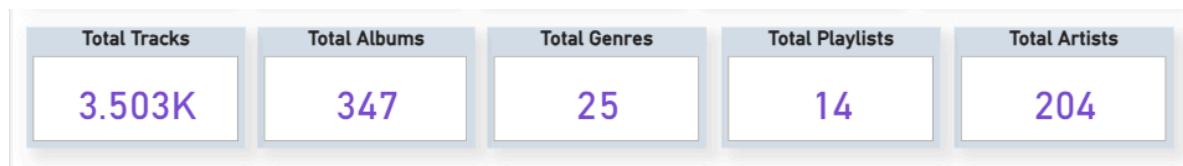
- CLV showed multiple peaks, particularly:
  - Mid-2017, early 2019, and mid-2020.
- The chart suggests consistent value from customers with minor fluctuations, indicating steady engagement over the years.

## PAGE 3 (MUSIC CONTENT INSIGHTS)



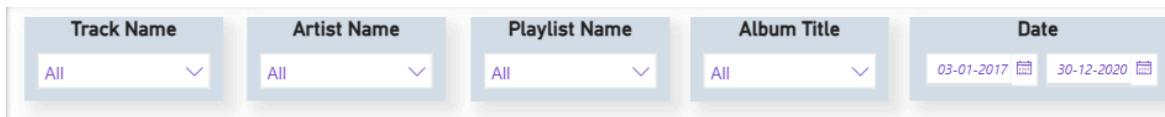
**Dataset used:** Music data merged (obtained by merging  
Album\_cleaned,Artist\_cleaned,Genre\_cleaned,Playlist\_cleaned,Playlist\_track\_cleaned,Track,Media\_type\_cleaned dataset)

### Key Metrics (Key Performance Indicators - KPIs):



- **Total Tracks:** 3.503K (3,503)
- **Total Albums:** 347
- **Total Genres:** 25
- **Total Playlists:** 14
- **Total Artists:** 204

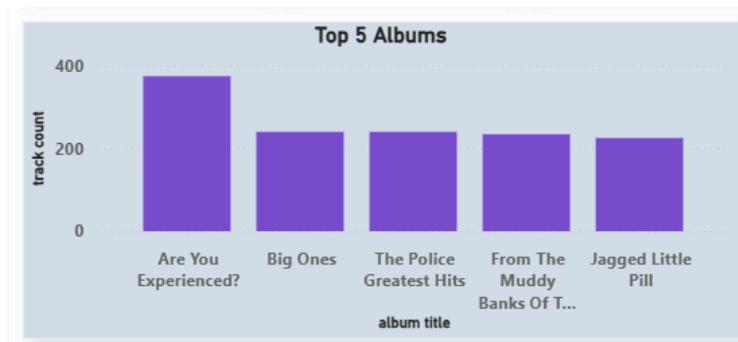
## Filters Available:



- **Track Name Filter:** "All"
- **Artist Name Filter** "All"
- **Playlist Name Filter:** "All"
- **Album Title Filter:** "All"
- **Date Range Filter:** "03-01-2017 - 30-12-2020"

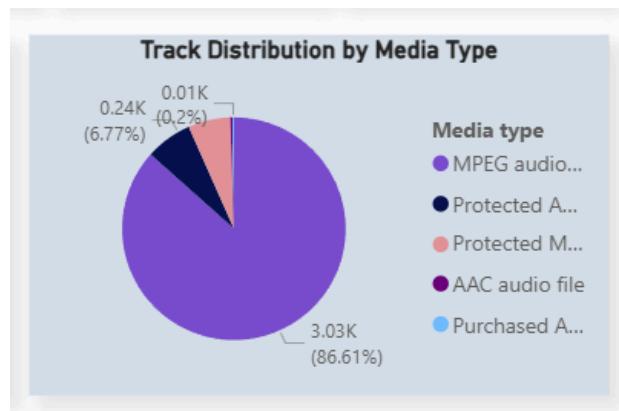
## Visual Insights:

### 1. Top 5 Albums (Bar Chart):



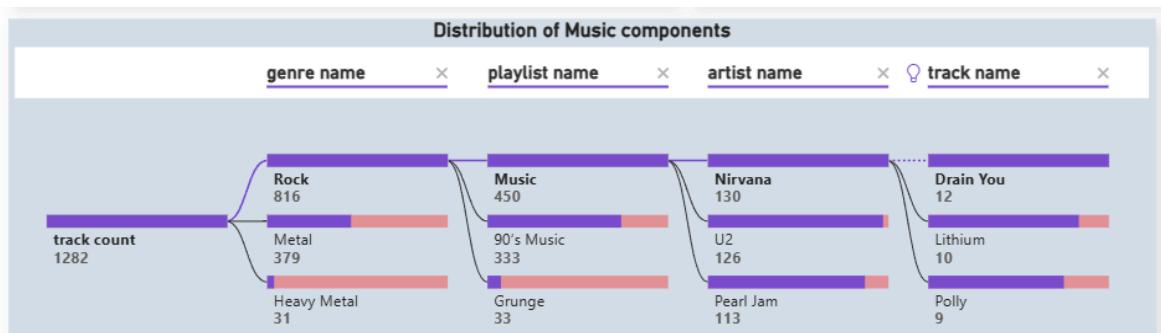
- "Are You Experienced?" is significantly the most represented album by track count.
- Rest albums show a clear dominance in terms of track presence range between 200-300.

## 2. Track Distribution by Media Type (Pie Chart):



- "MPEG audio file" is the overwhelmingly dominant media type, accounting for 86.61% of tracks.
- Other media types like "Protected AAC audio file", "Protected MPEG4 video file", "AAC audio file", and "Purchased AAC audio file" constitute a very small percentage of the track library.

## 3. Distribution of Music Components (Flow Chart):



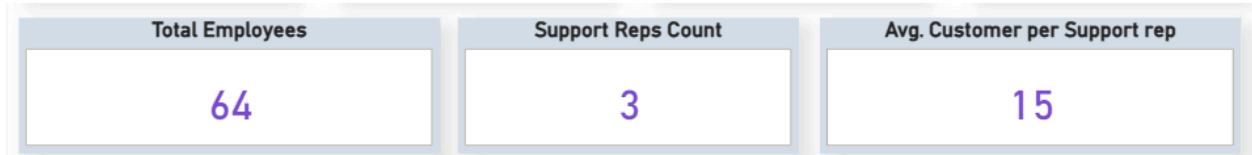
- **Genres:** "Rock" is the largest genre by track count (816), followed by "Metal" (379) and "Heavy Metal" (31).
- **Playlists:** "Music" is the largest playlist by track count (450), with "90's Music" (333) also being substantial.
- **Artists:** "Nirvana" has the most tracks (130), followed by "U2" (126) and "Pearl Jam" (113).
- **Tracks:** "Drain You" (12 tracks) and "Lithium" (10 tracks) are among the most frequently occurring individual tracks.

## PAGE 4 (EMPLOYEE & SUPPORT)



**Dataset used:** Employee Support Merged dataset (obtained by merging employee\_cleaned and customer\_cleaned dataset)

### Key Metrics (Key Performance Indicators - KPIs):



- **Total Employees:** 64
- **Support Representatives Count :** 3
- **Average customer per Support representative :** 15

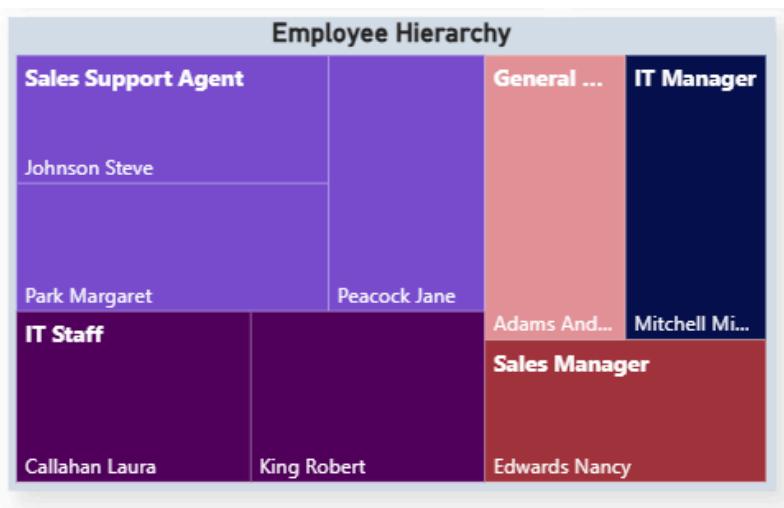
## Filters Available:

Employee ID	Employee Name	Employee Title	Date
All	All	All	01-05-2016  17-10-2017

- **Employee ID Filter:** "All"
- **Employee Name Filter** "All"
- **Employee Title Filter:** "All"
- **Date Range Filter:** "01-05-2016 - 17-10-2017"

## Visual Insights:

### 1. Employee Hierarchy (Treemap):



- The largest segments are "Sales Support Agent" and "IT Staff," indicating these are the largest departments or roles.
- Specific employees like "Johnson Steve" and "Park Margaret" are prominent within their respective departments, suggesting they might manage significant teams or responsibilities.
- "IT Manager" and "Sales Manager" are also defined, showing the management structure.

## 2. Customers Managed by Each Employee (Bar Chart):



- "Peacock Jane" manages the highest number of customers (around 20).
- "Park Margaret" and "Johnson Steve" also manage a significant number of customers, similar to "Peacock Jane."

## 3. Hire Dates Trend (Line Chart):



- There was a significant surge in employee hires around April 2017, dramatically increasing the total employee count.
- After the April 2017 spike, the employee count remained relatively stable until October 2017.

## Conclusion:

The four-page dashboard provides a complete view of revenue, customer behavior, music content, and employee performance. Revenue peaked in mid-2019 and late 2020, with the USA and Arizona being top contributors. František Wichterlova leads in customer spending, and customer lifetime value remains steady over time. Rock dominates the music library, with "MPEG audio file" being the main format. Most playlists are led by general music and 90s hits. Employee insights reveal strong support from reps like Peacock Jane, with a hiring surge in April 2017. Overall, the dashboard highlights steady business growth, global reach, and efficient employee support.